Proc. London Math. Soc. (3) 106 (2013) 819–858
doi:10.1112/plms/pds055

The influence of the first term of an arithmetic progression

Daniel Fiorilli

Abstract

The goal of this paper is to study the discrepancy of the distribution of arithmetic sequences in arithmetic progressions. We will fix a sequence $\mathcal{A} = \{\mathbf{a}(n)\}_{n \geqslant 1}$ of non-negative real numbers in a certain class of arithmetic sequences. For a fixed integer $a \neq 0$, we will be interested in the behaviour of $\mathcal{A}$ over the arithmetic progressions $a \bmod q$, on average over q. Our main result is that, for certain sequences of arithmetic interest, the value of a has a significant influence on this distribution, even after removing the first term of the progressions.

Contents

1. *Introduction*

The study of arithmetic sequences is a central problem in number theory. Undoubtedly, it is the sequence of prime numbers which has attracted the most attention among number theorists, leading to many theorems and conjectures. Other important sequences include integers which can be written as the sums of two squares, twin primes (primes p such that $p+2$ is also a prime), divisor sequences and so on. In general, number theorists are interested in sequences with arithmetical content, and one can formally define wide classes of such sequences. Some phenomena occurring in the theory of prime numbers happen to be true for much wider classes of arithmetic sequences, such as the Bombieri–Vinogradov theorem, for example (see [**3**, **4**, **5**]). Another example is the Granville–Soundararajan uncertainty principle (see [**11**]).

We will fix an integer $a \neq 0$ and study the distribution of an arithmetic sequence $\mathcal{A} = \{\mathbf{a}(n)\}_{n \geqslant 1}$ in the progressions $a \bmod q$, on average over q. The sequence $\mathcal{A}$ should be seen as the characteristic function of the set of integers having certain arithmetical properties, or a weighted version of such a function. For this reason, we will assume that the $\mathbf{a}(n)$ are real and non-negative. Under certain hypotheses, we will show how certain sequences remember the first term, that is, how the value of a can influence the distribution of $\mathcal{A}$ in the progressions $a \bmod q$. Examples of sequences covered by our analysis include the sequences of primes, sums of two squares (or more generally values of positive definite binary quadratic forms), prime k-tuples, that is, integers n such that $a_1 n + b_1, a_2 n + b_2, \dots, a_k n + b_k$ are all prime (conditionally) and integers without small prime factors. We will see that, in each of these examples, values of a which have the property that $\mathbf{a}(a) > 0$ have a negative influence. (By a having a negative

Received 26 April 2011; revised 4 April 2012; published online 31 October 2012.

2010 *Mathematics Subject Classification* 11B25, 11N13 (primary), 11N36, 11N56 (secondary).

influence, we mean that, on average over q, the residue classes a mod q contain fewer elements of the sequence $\mathcal{A}$ compared with other residue classes.) More mysteriously, there are other values of a having a negative influence, and it is not clear to me why these come up.

The structure of the paper is as follows. We begin in Section 2 by stating our concrete results for each of the arithmetic sequences mentioned earlier, to highlight the phenomena we will describe later on in more generality. In Section 3, we give a framework to study general arithmetic sequences and state the hypotheses on which our main theorems will depend. These hypotheses will be crucial in the proofs of Section 5. Our general results are stated in Section 4, and proved in Section 5. As we will see in Section 6, most of the concrete examples we give satisfy the hypotheses of Section 3, but in some cases we need to slightly modify the analysis. We also see in this section exactly which hypotheses are needed for each result.

2. *Examples*

Before we state the general result, let us look at concrete examples. Throughout, $\mathcal{A} = \{\mathbf{a}(n)\}_{n\geqslant 1}$ will be a fixed sequence of non-negative real numbers and $a \neq 0$ will be a fixed integer, on which every error term can possibly depend. We will adopt the convention that, for negative values of a, $\mathbf{a}(a) := 0$ (and similarly for $\Lambda(a)$). Moreover, $M = M(x)$ will denote a function tending to infinity with x. We define the following counting functions.

Definition 2.1.

$$\mathcal{A}(x) := \sum_{1\leqslant n\leqslant x} \mathbf{a}(n), \quad \mathcal{A}_d(x) := \sum_{\substack{1\leqslant n\leqslant x:\\ d\mid n}} \mathbf{a}(n), \quad \mathcal{A}(x;q,a) := \sum_{\substack{1\leqslant n\leqslant x\\ n\equiv a \bmod q}} \mathbf{a}(n).$$

2.1. *Primes*

The first example we give is the sequence of prime numbers.

Theorem 2.2. *Let $A > 0$ be a fixed real number and fix $\epsilon > 0$. We have, for $M = M(x) \leqslant (\log x)^A$, that*

$$\frac{1}{(\phi(a)/a)(x/M)} \sum_{\substack{q\leqslant x/M\\ (q,a)=1}} \left(\psi(x;q,a) - \Lambda(a) - \frac{\psi(x)}{\phi(q)}\right) \text{ is } \begin{cases} \sim -\dfrac{1}{2}\log M & \text{if } a = \pm 1,\\ \sim -\dfrac{1}{2}\log p & \text{if } a = \pm p^e,\\ = O(M^{-205/538+\epsilon}) & \text{otherwise,} \end{cases} \tag{1}$$

where the constant implied in O depends on a, ϵ and A.

Remark 2.3. This improves on a result of Friedlander and Granville [**10**], who showed that in this range of M, the left-hand side of (1) is $O(\log M)$. We refer the reader to [**9**] for a more detailed analysis of this case, as well as a more precise estimate.

2.2. *Integers represented by a fixed positive definite binary quadratic form, with multiplicity*

The second example we consider is the sequence of integers which can be represented by a fixed positive definite binary quadratic form $Q(x, y)$ with integer coefficients, counted with multiplicity, that is,

$$\mathbf{a}(n) := \#\{(x, y) \in \mathbb{Z}^2_{\geqslant 0} : Q(x, y) = n\}.$$

We will define $r_d(n)$ to be the total number of distinct representations of n by all of the inequivalent forms of discriminant d (which is not to be confused with $\mathbf{a}(n)$). By distinct representations, we mean that we count the representations modulo automorphisms of the forms. We also define the function

$$\rho_a(q) := \frac{1}{q} \cdot \#\{1 \leqslant x, y \leqslant q : Q(x,y) \equiv a \bmod q\}.$$

THEOREM 2.4. *Suppose that $Q(x,y) = \alpha x^2 + \beta xy + \gamma y^2$ is a fixed positive definite quadratic form (with integer coefficients) of discriminant $d := \beta^2 - 4\alpha\gamma < 0$, with $(\alpha, \beta, \gamma) = 1$ and $d \equiv 1, 5, 9, 12, 13 \bmod 16$ (for simplicity). Fix an integer a such that $(a, 2d) = 1$ and fix $\epsilon > 0$. We have, for $M = M(x) \leqslant x^\lambda$ where $\lambda < \frac{1}{12}$ is a fixed real number, that*

$$\frac{1}{x/M} \sum_{q \leqslant x/M} \left(\mathcal{A}(x; q, a) - \mathbf{a}(a) - \frac{\rho_a(q)}{q} \mathcal{A}(x) \right) = -C_Q \rho_a(4d) r_d(|a|) + O\left(\frac{1}{M^{1/3-\epsilon}} \right), \tag{2}$$

with

$$C_Q := \frac{A_Q}{2L(1, \chi_d)} \left(= \frac{w_d \sqrt{|d|}}{4\pi h_d} A_Q \right),$$

where A_Q is the area of the region $\{(x,y) \in \mathbb{R}^2_{\geqslant 0} : Q(x,y) \leqslant 1\}$, $\chi_d := (4d/\cdot)$, w_d is the number of units of $\mathbb{Q}(\sqrt{d})$ and h_d is its class number. The constant implied in O depends on a, ϵ, λ and Q.

REMARK 2.5. The number $\rho_a(4d)$ is either zero or equal to $2^{\omega(2d)}$, $2^{\omega(2d)-2}$ or $3 \cdot 2^{\omega(2d)-2}$, depending on $Q(x,y)$ ($\omega(n)$, denotes the number of distinct prime factors of n) (see Lemma A.3). For this reason, if $\rho_a(4d) > 0$, then it is independent of a.

Therefore, there is no bias if $\rho_a(4d) = 0$ or if $|a|$ cannot be represented by a form of discriminant d. However, if this is not the case, then the bias is proportional to the number of such representations.

2.3. *Sums of two squares, without multiplicity*

The next example is the sequence of integers which can be written as the sum of two squares, without multiplicity. We define

$$\mathbf{a}(n) := \begin{cases} 1 & \text{if } n = \square + \square, \\ 0 & \text{else.} \end{cases}$$

For a fixed odd integer a, we define the multiplicative function $\mathbf{g}_a(q)$ on prime powers as follows. For $p \neq 2$ such that $p^f \parallel a$ with $f \geqslant 0$,

$$\mathbf{g}_a(p^e) := \frac{1}{p^e} \times \begin{cases} 1 & \text{if } p \equiv 1 \bmod 4, \\ 1 & \text{if } p \equiv 3 \bmod 4, e \leqslant f, 2 \mid e, \\ \dfrac{1}{p} & \text{if } p \equiv 3 \bmod 4, e \leqslant f, 2 \nmid e, \\ 1 + \dfrac{1}{p} & \text{if } p \equiv 3 \bmod 4, e > f, 2 \mid f, \\ 0 & \text{if } p \equiv 3 \bmod 4, e > f, 2 \nmid f. \end{cases} \tag{3}$$

Moreover, $\mathbf{g}_a(2) := \frac{1}{2}$ and, for $e \geqslant 2$, $\mathbf{g}_a(2^e) := (1 + (-1)^{(a-1)/2})/2^{e+2}$.

THEOREM 2.6. *Fix an integer $a \equiv 1 \bmod 4$. We have, for $M = M(x) \leqslant (\log x)^{\lambda}$ where $\lambda < \frac{1}{5}$ is a fixed real number, that*

$$\frac{1}{x/2M} \sum_{x/2M<q\leqslant x/M} (\mathcal{A}(x;q,a) - \mathbf{a}(a) - \mathbf{g}_a(q)\mathcal{A}(x)) \\ \sim -\left(\frac{\log M}{\log x}\right)^{1/2} \frac{(-4)^{-\ell_a-1}(2\ell_a+2)!}{(4\ell_a^2-1)(\ell_a+1)!\pi} \prod_{\substack{p^f \| a: \\ p\equiv 3 \bmod 4, \\ f odd}} \frac{\log(p^{(f+1)/2})}{\log M}, \tag{4}$$

where $\ell_a := \#\{p^f \parallel a : p \equiv 3 \bmod 4, 2 \nmid f\}$ is the number of primes dividing a to an odd power and which are congruent to 3 mod 4.

REMARK 2.7. The right-hand side of (4) is $o((\log x)^{-1/2})$ if and only if $|a|$ cannot be written as the sum of two squares. Also, if $|a| = \square + \square$, then it is equal to $-(1/2\pi)(\log M/\log x)^{1/2}$. Moreover, one can show that if $a \equiv 3 \bmod 4$, then the left-hand side of (4) is always $o((\log x)^{-1/2})$.

2.4. *Prime k-tuples*

The next example concerns prime k-tuples. Let $\mathcal{H} = \{\mathcal{L}_1, \dots, \mathcal{L}_k\}$ be a k-tuple of distinct linear forms $\mathcal{L}_i(n) = a_i n + b_i$, with $a_i, b_i \in \mathbb{Z}$, $a_i \geqslant 1$, and define

$$\mathcal{P}(n;\mathcal{H}) := \prod_{\mathcal{L}\in\mathcal{H}} \mathcal{L}(n).$$

We will suppose that $\mathcal{H}$ is admissible, that is, for every prime p,

$$\nu_{\mathcal{H}}(p) := \#\{x \bmod p : \mathcal{P}(x;\mathcal{H}) \equiv 0 \bmod p\} < p.$$

Define

$$\mathbf{a}(n) := \prod_{\mathcal{L}\in\mathcal{H}} \Lambda(\mathcal{L}(n)) = \Lambda(a_1 n + b_1)\Lambda(a_2 n + b_2)\cdots\Lambda(a_k n + b_k).$$

The singular series associated to $\mathcal{H}$ is

$$\mathfrak{S}(\mathcal{H}) := \prod_p \left(1 - \frac{\nu_{\mathcal{H}}(p)}{p}\right)\left(1 - \frac{1}{p}\right)^{-k}.$$

Note that if $(\mathcal{P}(a;\mathcal{H}), q) > 1$, then $\mathcal{A}(x;q,a)$ is bounded. Fix $\delta > 0$. The Hardy–Littlewood conjecture stipulates that there exists a function $\mathbf{L}(x)$ tending to infinity with x such that if $(\mathcal{P}(a;\mathcal{H}), q) = 1$,

$$\mathcal{A}(x) = \mathfrak{S}(\mathcal{H})x + O\left(\frac{x}{\mathbf{L}(x)^{2+2\delta}}\right). \tag{5}$$

Define

$$\gamma(q) := \prod_{p|q} \left(1 - \frac{\nu_{\mathcal{H}}(p)}{p}\right).$$

THEOREM 2.8. *Assume that (5) holds uniformly for all admissible k-tuples $\tilde{\mathcal{H}}$ such that $|a_i| \leqslant \mathbf{L}(x)^{1+\delta}$ and $|b_i| = O(1)$. Fix a k-tuple $\mathcal{H} = \{\mathcal{L}_1, \dots, \mathcal{L}_k\}$. We have, for*

$M = M(x) \leqslant \mathbf{L}(x)$, *that the average*

$$\frac{1}{(\phi(\mathcal{P}(a;\mathcal{H}))/\mathcal{P}(a;\mathcal{H}))(x/2M)} \sum_{\substack{x/2M<q\leqslant x/M:\\ (q,\mathcal{P}(a;\mathcal{H}))=1}} \left(\mathcal{A}(x;q,a) - \mathbf{a}(a) - \frac{\mathcal{A}(x)}{q\gamma(q)}\right) \text{ is}$$

$$\begin{cases} \sim -\dfrac{(\log M)^{k-\omega(P(a;\mathcal{H}))}}{2(k-\omega(\mathcal{P}(a;\mathcal{H})))!} \displaystyle\prod_{p|\mathcal{P}(a;\mathcal{H})} \frac{p-\nu_{\mathcal{H}}(p)}{p-1} \log p & \text{if } \omega(\mathcal{P}(a;\mathcal{H})) \leqslant k, \\ = O(M^{-\delta_k}) & \text{otherwise,} \end{cases}$$

where $\delta_k > 0$ *is a positive real number depending on* k, *and* $\omega(n)$ *denotes the number of distinct prime factors of* n. *The constant implied in* O *depends on* a, δ *and* $\mathcal{H}$.

In the case of twin primes, we have $\mathcal{H} = \{n, n+2\}$, so $\mathcal{P}(a,\mathcal{H}) = a(a+2)$, and the function $\nu_{\mathcal{H}}$ is given by $\nu_{\mathcal{H}}(2) = 1$ and $\nu_{\mathcal{H}}(p) = 2$ for odd p. We obtain that the average is

$$\begin{cases} \sim -(\log M)^2/4 & \text{if } a = -1, \\ \sim -\dfrac{\log 3}{4} \log M & \text{if } a = 1, -3, \\ \sim -\dfrac{\log 2}{2} \log M & \text{if } a = 2, -4, \\ \sim -\dfrac{\log p \log q}{2} \dfrac{p-\nu_{\mathcal{H}}(p)}{p-1} \dfrac{q-\nu_{\mathcal{H}}(q)}{q-1} & \text{if } a(a+2) = \pm p^e q^f, \\ O(M^{-\delta_2}) & \text{if } \omega(a(a+2)) \geqslant 3. \end{cases}$$

2.5. *Integers free of small prime factors*

For $y = y(x)$ a function of x, we define

$$\mathbf{a}_y(n) := \begin{cases} 1 & \text{if } p \mid n \Rightarrow p \geqslant y \\ 0 & \text{else} \end{cases} \qquad \mathcal{A}(x,y) := \sum_{n\leqslant x} \mathbf{a}_y(n),$$

$$\gamma_y(q) := \prod_{\substack{p|q\\ p<y}} \left(1 - \frac{1}{p}\right), \quad \mathcal{A}(x,y;q,a) := \sum_{\substack{n\leqslant x\\ n\equiv a \bmod q}} \mathbf{a}_y(n).$$

THEOREM 2.9. *Fix* $a \neq 0$, $0 < \delta < \frac{1}{2}$ *and* $M = M(x) \leqslant (\log x)^{1-\delta}$. *If*

$$\nu_y(a,M) := \frac{1}{(x/2M)(\phi(a)/a)} \sum_{\substack{x/2M<q\leqslant x/M\\ (q,a)=1}} \left(\mathcal{A}(x,y;q,a) - \mathbf{a}_y(a) - \frac{\mathcal{A}(x,y)}{q\gamma_y(q)}\right),$$

then, for $y \leqslant e^{(\log M)^{1/2-\delta}}$ *with* $y \to \infty$,

$$\nu_y(a,M) = \begin{cases} -\frac{1}{2} + o(1) & \text{if } a = \pm 1, \\ o(1) & \text{otherwise,} \end{cases}$$

and, for $(\log x)^{\log\log\log x} \leqslant y \leqslant \sqrt{x}$,

$$\nu_y(a,M) = \frac{\mathcal{A}(x,y)}{x} \times \begin{cases} \left(-\dfrac{1}{2} + o(1)\right) \log M & \text{if } a = \pm 1, \\ -\dfrac{1}{2} \log p + o(1) & \text{if } a = \pm p^k, \\ o(1) & \text{otherwise.} \end{cases}$$

(*We have no result in the intermediate range.*)

Remark 2.10. For x large enough, $\mathbf{a}_y(a) = 0$ unless $a = \pm 1$.

3. *Definitions and hypotheses*

3.1. *Arithmetic sequences*

The goal of this section is to give a framework to study a class of arithmetic sequences, which will justify the hypotheses of the next section. This discussion is modelled on that in [11].

We wish to study the sequence $\mathcal{A} = \{\mathbf{a}(n)\}_{n\geqslant 1}$ in arithmetic progressions, therefore one of our goals will be to prove the existence of a multiplicative function $\mathbf{g}_a(q)$ such that

$$\mathcal{A}(x; q, a) \sim \mathbf{g}_a(q)\mathcal{A}(x),$$

whenever $\mathbf{g}_a(q) \neq 0$. Let us give a heuristic way to do this with the help of an auxiliary multiplicative function $\mathbf{h}(d)$. First, denote by $\mathcal{S}$ a finite set of 'bad primes', which are inherent to the sequence $\mathcal{A}$. We will assume that $\mathcal{A}$ is well distributed in the progressions 0 mod d, that is, there exists a multiplicative function $\mathbf{h}(d)$ such that, for $(d, \mathcal{S}) = 1$,

$$\mathcal{A}_d(x) \approx \frac{\mathbf{h}(d)}{d}\mathcal{A}(x).$$

Moreover, we require that, for each prime p, $0 \leqslant \mathbf{h}(p) < p$. As was remarked in [11], this is a very mild assumption, and it holds for each of the sequences we are interested in. The fact that $\mathbf{h}(d)$ is multiplicative can be rephrased as 'the events that $\mathbf{a}(n)$ is divisible by coprime integers are independent'. Let us also assume that

$$\mathcal{A}(x; q, a) \approx \frac{1}{\phi(q/(q,a))} \sum_{\substack{n\leqslant x:\\ (q,n)=(q,a)}} \mathbf{a}(n),$$

that is, the sum is equally partitioned among the $\phi(q/(q, a))$ arithmetic progressions b mod q with $(b, q) = (a, q)$. We then compute

$$\begin{aligned}
\mathcal{A}(x; q, a) &\approx \frac{1}{\phi(q/(q,a))} \sum_{\substack{n\leqslant x\\ (q,n)=(q,a)}} \mathbf{a}(n)\\
&= \frac{1}{\phi(q/(q,a))} \sum_{d\mid q/(q,a)} \mu(d)\mathcal{A}_{(q,a)d}(x)\\
&\approx \mathcal{A}(x)\frac{1}{\phi(q/(q,a))} \sum_{d\mid q/(q,a)} \mu(d)\frac{\mathbf{h}((q,a)d)}{(q,a)d}\\
&= \mathbf{g}_a(q)\mathcal{A}(x),
\end{aligned}$$

where

$$\mathbf{g}_a(q) = \mathbf{g}_{(a,q)}(q) := \frac{1}{\phi(q/(q,a))} \sum_{d\mid q/(q,a)} \mu(d)\frac{\mathbf{h}((q,a)d)}{(q,a)d}$$

is a multiplicative function of q which depends on (q, a) (rather than depending on a). We have thus expressed the multiplicative function $\mathbf{g}_a(q)$ in terms of $\mathbf{h}(d)$. More explicitly, we have, when $p^f \parallel a$ (with $(pa, \mathcal{S}) = 1$), that

$$\mathbf{g}_a(p^e) = \begin{cases} \dfrac{\mathbf{h}(p^e)}{p^e} & \text{if } e \leqslant f,\\[2ex] \dfrac{1}{\phi(p^e)}\left(\mathbf{h}(p^f) - \dfrac{\mathbf{h}(p^{f+1})}{p}\right) & \text{if } e > f. \end{cases} \tag{6}$$

In particular, if $p \nmid a$,

$$\mathbf{g}_a(p^e) = \frac{1}{\phi(p^e)}\left(1 - \frac{\mathbf{h}(p)}{p}\right).$$

Another way to write this is

$$\mathcal{A}(x; q, a) \approx \frac{\mathbf{f}_a(q)}{q\gamma(q)}\mathcal{A}(x), \tag{7}$$

where

$$\gamma(q) := \frac{\phi(q)}{q}\prod_{p|q}\left(1 - \frac{\mathbf{h}(p)}{p}\right)^{-1} = \prod_{p|q}\frac{1-1/p}{1-\mathbf{h}(p)/p},$$

and $\mathbf{f}_a(q)$ is a multiplicative function defined by $\mathbf{f}_a(q) := \mathbf{g}_a(q)q\gamma(q)$. Note that for $(a, q) = 1$, $\mathbf{f}_a(q) = 1$.

3.2. *Hypotheses*

In the following, $\delta > 0$ will denote a (small) fixed real number which will change from one statement to another. We will also fix an integer $a \neq 0$ with the property that $(a, \mathcal{S}) = 1$, where $\mathcal{S}$ is a finite set of bad primes. The function $\mathbf{L} : [0, \infty) \to [1, \infty)$ will be a given positive increasing function such that $\mathbf{L}(x) \to \infty$ as $x \to \infty$ (think of $\mathbf{L}(x)$ as a power of $\log x$). We now assume the existence of a multiplicative function $\mathbf{f}_a(q) = \mathbf{f}_{(a,q)}(q)$, depending on (a, q), and of $\gamma(q) \neq 0$, which is independent of a (as in Section 3.3), such that, for any fixed $a \neq 0$ and $q \geqslant 1$,

$$\mathcal{A}(x; q, a) \sim \frac{\mathbf{f}_a(q)}{q\gamma(q)}\mathcal{A}(x),$$

whenever $\mathbf{f}_a(q) \neq 0$. To simplify the notation, we will also assume the existence of a multiplicative function $\mathbf{h}(d)$ such that (6) holds (for $(qa, \mathcal{S}) = 1$), and we define

$$\mathbf{g}_a(q) := \frac{\mathbf{f}_a(q)}{q\gamma(q)}.$$

The existence of such functions is justified by the heuristic argument of the last section.

Hypothesis 3.1. There exists $\delta > 0$ and a positive increasing function $\mathbf{R}(x)$ (think of $\mathbf{R}(x)$ as a small power of x), with $\mathbf{L}(x)^{1+\delta} \leqslant \mathbf{R}(x) \leqslant \sqrt{x}$, such that

$$\sum_{q \leqslant 2\mathbf{R}(x)} \max_{y \leqslant x} |\mathcal{A}(y; q, a) - \mathbf{g}_a(q)\mathcal{A}(y)| \ll \frac{\mathcal{A}(x)}{\mathbf{L}(x)^{1+\delta}}.$$

We will see later that if we use dyadic intervals, we can replace Hypothesis 3.1 by a weaker hypothesis.

Hypothesis 3.1*. There exists $\delta > 0$ such that

$$\sum_{q \leqslant 2\mathbf{L}(x)} \max_{y \leqslant x} |\mathcal{A}(y; q, a) - \mathbf{g}_a(q)\mathcal{A}(y)| \ll \frac{\mathcal{A}(x)}{\mathbf{L}(x)^{1+\delta}}.$$

Hypothesis 3.2. There exists $\delta > 0$ such that, for any $z = z(x)$ in the range $1/\mathbf{L}(x) \leqslant z(x) \leqslant 1 + |a|/x$, we have

$$\frac{\mathcal{A}(zx)}{\mathcal{A}(x)} = z + O\left(\frac{1}{\mathbf{L}(x)^{1+\delta}}\right).$$

Moreover, for $n \leqslant x$, we have the following bound:

$$\mathbf{a}(n) \ll \frac{\mathcal{A}(x)}{\mathbf{L}(x)^{1+\delta}}.$$

The next hypothesis is somewhat more specific to our analysis than the ones above, and it will allow us to use the analytic theory of zeta functions.

HYPOTHESIS 3.3. There exists a real number $\mathbf{k} \geqslant 0$ such that the sum

$$\sum_{p \notin \mathcal{S}} \frac{\mathbf{h}(p) - \mathbf{k}}{p}$$

is convergent. More generally, there exists $\delta > 0$ such that, for any real number t and integer $n \geqslant 1$, we have, for any $\epsilon > 0$,

$$\sum_{\substack{p \leqslant x \\ p \notin \mathcal{S}}} \frac{\mathbf{h}(p) - \mathbf{k}}{p^{1+it}} \leqslant \left(\frac{1}{2} - \delta\right) \log(|t| + 2) + O(1),$$

$$\sum_{\substack{p \leqslant x \\ p \notin \mathcal{S}}} \frac{(\mathbf{h}(p) - \mathbf{k}) \log^n p}{p^{1+it}} \ll_{n,\epsilon} (|t| + 2)^{\epsilon}.$$

Finally, $\mathbf{h}(p) < p$ and, for any $\epsilon > 0$,

$$\mathbf{h}(d) \ll_{\epsilon} d^{\epsilon}.$$

The final hypothesis will be useful when studying the full interval $1 \leqslant q \leqslant x/M$ rather than a dyadic one. It is not known whether this hypothesis holds for all sequences considered in Section 2; for this reason we used dyadic intervals in Theorems 2.6, 2.8 and 2.9.

HYPOTHESIS 3.4. There exists $\delta > 0$ such that, with the same $\mathbf{R}(x)$ as in Hypothesis 3.1, we have

$$\sum_{q \leqslant x/\mathbf{R}(x)} (\mathcal{A}^*(x; q, a) - \mathbf{g}_a(q)\mathcal{A}(x)) \ll \frac{\mathcal{A}(x)}{\mathbf{L}(x)^{1+\delta}},$$

where $\mathcal{A}^*(x; q, a)$ is defined as in (26).

3.3. *The formula for the average*

In this section, we give a formula for the 'average' $\mu_{\mathbf{k}}(a, M)$ that will appear in Theorems 4.1 and 4.1*. The formula is rather complicated in its general form; however, in concrete examples it can be seen that it reflects the nature of the sequence $\mathcal{A}$.

DEFINITION 3.5.

$$\omega_{\mathbf{h}}(a) := \#\{p^f \parallel a \text{ with } f \geqslant 1 : \mathbf{h}(p^f) = \mathbf{h}(p^{f+1})/p\}.$$

Definition 3.6. Assume Hypothesis 3.3 and suppose that $\mathcal{S} = \emptyset$. For an integer $a \neq 0$ and a real number $\mathbf{k} \geqslant 0$, we define

$$\mu_{\mathbf{k}}(a, M) := -\frac{1}{2}\frac{(\log M)^{1-\mathbf{k}-\omega_{\mathbf{h}}(a)}}{\Gamma(2-\mathbf{k}-\omega_{\mathbf{h}}(a))} \prod_{\substack{p^f \| a: \\ \mathbf{h}(p^f)=\mathbf{h}(p^{f+1})/p, \\ f\geqslant 0}} \frac{1+\mathbf{h}(p)+\cdots+\mathbf{h}(p^f)}{(1-1/p)^{\mathbf{k}-1}}\log p$$

$$\times \prod_{\substack{p^f \| a: \\ \mathbf{h}(p^f)\neq\mathbf{h}(p^{f+1})/p, \\ f\geqslant 0}} \frac{\mathbf{h}(p^f)-\mathbf{h}(p^{f+1})/p}{(1-1/p)^{\mathbf{k}}}. \tag{8}$$

Remark 3.7. The two products appearing in (8) can potentially be infinite, since we are allowing f to be zero. The first of these products, however, is a finite product, since, for all but a finite number of primes, we have $f = 0$, and $\mathbf{h}(p) < p$ for all p, so $\mathbf{h}(1) \neq \mathbf{h}(p)/p$. The second product is convergent, since, for $p \nmid a$, we have $\mathbf{h}(p^f) - \mathbf{h}(p^{f+1})/p = 1 - \mathbf{h}(p)/p \approx 1 - \mathbf{k}/p$. Of course both these statements rely on the assumption of Hypothesis 3.3.

Remark 3.8. One sees that, for integer values of $\mathbf{k}$, $\mu_{\mathbf{k}}(a, M) = 0$ if and if only $\omega_{\mathbf{h}}(a) \geqslant 2 - \mathbf{k}$, by the location of the poles of $\Gamma(s)$. Moreover, since these are the only poles, we have $\mu_{\mathbf{k}}(a, M) \neq 0$ whenever $\mathbf{k} \notin \mathbb{Z}$.

Remark 3.9. If $\mathcal{S} \neq \emptyset$, we can still give a formula for $\mu_{\mathbf{k}}(a, M)$, assuming we understand well $\mathbf{g}_a(p^e)$ with $p \in \mathcal{S}$. However, this would complicate the already lengthy definition of $\mu_{\mathbf{k}}(a, M)$, so we only give individual descriptions in the examples. Two instances of $\mathcal{S} \neq \emptyset$ are given in the proofs of Theorems 2.4 and 2.6.

4. *Main result*

The main result of the paper is a formula for the average value of the discrepancy $\mathcal{A}(x; q, a) - \mathbf{g}_a(q)\mathcal{A}(x)$, summed over $1 \leqslant q \leqslant Q$, with Q large enough in terms of x.

Theorem 4.1. *Assume that Hypotheses 3.1–3.4 hold with $\mathcal{S} = \emptyset$ and the function $\mathbf{L}(x)$. Fix an integer $a \neq 0$ and let $M = M(x)$ be a function tending to infinity with x such that $M(x) \leqslant \mathbf{L}(x)$. We have, for any fixed real number $A > 0$, that*

$$\sum_{q\leqslant x/M} (\mathcal{A}(x; q, a) - \mathbf{a}(a) - \mathbf{g}_a(q)\mathcal{A}(x)) = \frac{\mathcal{A}(x)}{M}\left(\mu_{\mathbf{k}}(a, M)(1 + o(1)) + O\left(\frac{1}{\log^A M}\right)\right), \tag{9}$$

where $\mathbf{a}(a)$ is the first term of $\mathcal{A}(x; q, a)$ for positive a, and, whenever a is negative, we set $\mathbf{a}(a) = 0$. The constant implied in O depends on a, A and $\mathcal{A}$.

We also give a dyadic version, which assumes a weaker form of Hypothesis 3.1, and does not assume Hypothesis 3.4 at all.

Theorem 4.1*. *Assume that Hypotheses 3.1*, 3.2 and 3.3 hold with $\mathcal{S} = \emptyset$ and the function $\mathbf{L}(x)$. Fix an integer $a \neq 0$ and let $M = M(x)$ be a function tending to infinity with*

x such that $M(x) \leqslant \mathbf{L}(x)$. We have, for any fixed real number $A > 0$, that

$$\sum_{x/2M<q\leqslant x/M} (\mathcal{A}(x;q,a) - \mathbf{a}(a) - \mathbf{g}_a(q)\mathcal{A}(x)) = \frac{\mathcal{A}(x)}{2M}\left(\mu_{\mathbf{k}}(a,M)(1+o(1)) + O\left(\frac{1}{\log^A M}\right)\right). \tag{10}$$

The constant implied in O depends on a, A and $\mathcal{A}$.

Remark 4.2. As we have seen in the examples of Section 2, Theorems 4.1 and 4.1* easily generalize to arbitrary (given) sets $\mathcal{S} \neq \emptyset$, as long as we understand $\mathbf{g}_a(p^e)$ for each $p \in \mathcal{S}$. In the general case, one needs to adapt the lemmas of Section 4. One should add a product over the primes in $\mathcal{S}$ to $\mathfrak{S}_2(s)$ (and remove these primes from the product defining $Z_3(s)$). If we assume that $f_a(p^e)$ stabilizes for e large enough, that is, there exists e_0 such that $f_a(p^e) = f_a(p^{e_0})$ for all $e > e_0$, then we can give a general formula for $\mu_{\mathbf{k}}(a,M)$. Note that, depending on $g_a(p^e)$ for $e \leqslant e_0$, the extra factors of $\mathfrak{S}_2(s)$ for primes $p \in \mathcal{S}$ can possibly vanish at $s = -1$ and thus change the analytic behaviour of $Z(s)$, ending up in either changing $\mu_{\mathbf{k}}(a,M)$ by a factor of $(\log M)^{-1}$ or making it vanish.

Remark 4.3. If $\mu_{\mathbf{k}}(a,M) \neq 0$, then Theorems 4.1 and 4.1* give asymptotics for the sum on the left-hand side.

Remark 4.4. Suppose that $\mathbf{k} = 0$ (for example, when $\mathcal{A}$ is the sequence of primes).
If $\omega_{\mathbf{h}}(a) \geqslant 2$, then $\mu_0(a,M) = 0$.
If $\omega_{\mathbf{h}}(a) = 1$, so there is a unique $p_0^{f_0} \parallel a$, $f_0 \geqslant 1$, such that $\mathbf{h}(p_0^{f_0}) = \mathbf{h}(p_0^{f_0+1})/p_0$, then

$$\mu_0(a,M) = -\frac{1}{2}\left(1 - \frac{1}{p_0}\right)(1 + \mathbf{h}(p_0) + \cdots + \mathbf{h}(p_0^{f_0}))\log p_0 \prod_{\substack{p^f \parallel a \\ f \geqslant 0 \\ p \neq p_0}} (\mathbf{h}(p^f) - \mathbf{h}(p^{f+1})/p).$$

If $\omega_{\mathbf{h}}(a) = 0$, then

$$\mu_0(a,M) = -\frac{\log M}{2} \prod_{\substack{p^f \parallel a \\ f \geqslant 0}} (\mathbf{h}(p^f) - \mathbf{h}(p^{f+1})/p).$$

Remark 4.5. Suppose that $\mathbf{k} = 1$ (for example, when $\mathcal{A}$ is the sequence of integers that can be written as the sum of two squares, counted with multiplicity). Then

$$\mu_1(a,M) = -\frac{1}{2} \prod_{\substack{p^f \parallel a \\ f \geqslant 0}} \frac{\mathbf{h}(p^f) - \mathbf{h}(p^{f+1})/p}{1 - 1/p}.$$

Remark 4.6. Suppose that $\mathbf{k}$ is an integer $\geqslant 2$ (for example, when $\mathcal{A}$ is the sequence of integers of the form $(m + c_1)(m + c_2)\cdots(m + c_{\mathbf{k}})$, where the c_i are distinct integers). Then $\mu_1(a,M) = 0$.

5. *Proof of the main result*

The goal of this section is to prove Theorems 4.1 and 4.1*.

5.1. *An estimate for the main sum*

In this section, we will assume that $\mathcal{S} = \emptyset$ for simplicity.

PROPOSITION 5.1. *Assume Hypothesis* 3.3. *Let* $M = M(x)$ *and* $\mathbf{R} = \mathbf{R}(x)$ *be two positive functions of* x *such that* $M(x)^{1+\delta} \leqslant \mathbf{R}(x) \leqslant \sqrt{x}$ *for a fixed* $\delta > 0$. *We have*

$$\sum_{1\leqslant r\leqslant \mathbf{R}} \mathbf{g}_a(r)\left(1-\frac{r}{\mathbf{R}}\right) - \sum_{1\leqslant r\leqslant M} \mathbf{g}_a(r)\left(1-\frac{r}{M}\right) - \sum_{x/\mathbf{R}<q\leqslant x/M} \mathbf{g}_a(q)$$
$$= \frac{\mu_k(a,M)}{M}\left(1+O\left(\frac{\log\log M}{\log M}\right)\right) + O_{A,\delta}\left(\frac{1}{M\log^A M}\right).$$

The proof of Proposition 5.1 will require several lemmas.

LEMMA 5.2. *With* $\mathbf{f}_a(n)$ *and* $\gamma(n)$ *defined as in Section* 3.1, *we have*

$$\mathbf{g}_a(n) \ll \frac{1}{\phi(n)}.$$

Proof. By definition,

$$\begin{aligned}\mathbf{g}_a(n) &= \prod_{p^e\|n} \mathbf{g}_a(p^e) \ll_{a,\mathcal{S}} \prod_{\substack{p^e\|n\\ p\nmid a, p\notin\mathcal{S}}} \mathbf{g}_a(p^e)\\ &= \prod_{\substack{p^e\|n\\ p\nmid a, p\notin\mathcal{S}}} \frac{1}{\phi(p^e)}\left(1-\frac{\mathbf{h}(p)}{p}\right)\\ &\leqslant \prod_{\substack{p^e\|n\\ p\nmid a, p\notin\mathcal{S}}} \frac{1}{\phi(p^e)} \ll_{a,\mathcal{S}} \frac{1}{\phi(n)}.\end{aligned}$$

since $\mathbf{h}(p) \geqslant 0$ (The proof also works when $\mathcal{S}$ is nonempty.) $\square$

LEMMA 5.3. *Assume Hypothesis* 3.3. *Let* $\eta : [0,\infty) \to [0,\infty)$ *be a piecewise continuous function supported on* $[0,1]$, *taking a value halfway between the limit values at discontinuities, and suppose the integral*

$$\mathcal{M}\eta(s) := \int_0^1 \eta(x)x^{s-1}\,dx$$

converges absolutely for $\Re(s) > 0$. *Then*

$$\sum_{n\leqslant M} \mathbf{g}_a(n)\eta\left(\frac{n}{M}\right) = \frac{1}{2\pi i}\int_{\Re(s)=1} \mathfrak{S}_2(s)\zeta(s+1)\zeta(s+2)^{1-\mathbf{k}} Z_3(s)\mathcal{M}\eta(s)M^s\,ds, \tag{11}$$

where

$$\begin{aligned}\mathfrak{S}_2(s) = \mathfrak{S}_2(s;a) &:= \prod_{\substack{p^f\|a\\ f\geqslant 1}} \left[\left(1+\frac{\mathbf{h}(p)}{p^{s+1}}+\cdots+\frac{\mathbf{h}(p^f)}{p^{f(s+1)}}\right)\left(1-\frac{1}{p^{s+1}}\right)\right.\\ &\quad \left.+\frac{\mathbf{h}(p^f)-\mathbf{h}(p^{f+1})/p}{1-1/p}\frac{1}{p^{(f+1)(s+1)}}\right]\left(1-\frac{1}{p^{s+2}}\right)^{1-\mathbf{k}},\\ Z_3(s) = Z_3(s;a) &:= \prod_{p\nmid a}\left(1+\frac{1}{p^{s+1}}\left(\frac{1}{\gamma(p)}-1\right)\right)\left(1-\frac{1}{p^{s+2}}\right)^{1-\mathbf{k}}.\end{aligned} \tag{12}$$

Moreover, $\mathfrak{S}_2(s)$ *is holomorphic for* $\Re(s) > -2$ *and* $Z_3(s)$ *is holomorphic for* $\Re(s) > -1$.

Proof. Define

$$Z_{\mathcal{A}}(s) = Z_{\mathcal{A}}(s;a) := \sum_{n=1}^{\infty} \frac{\mathbf{g}_a(n)}{n^s} = \prod_p \left(1 + \frac{\mathbf{g}_a(p)}{p^s} + \frac{\mathbf{g}_a(p^2)}{p^{2s}} + \cdots\right).$$

A standard computation using the definition of $\mathbf{g}_a(n)$ (see (6)) yields that

$$Z_{\mathcal{A}}(s) = \mathfrak{S}_2(s)\zeta(s+1)\zeta(s+2)^{1-\mathbf{k}} Z_3(s).$$

If $\mathbf{k} \notin \mathbb{Z}$, the function $\mathfrak{S}_2(s)$ has branching points at $s = -2 + 2\pi i\ell/\log p$ for each $\ell \in \mathbb{Z}$ and $p \mid a$. Restricting ourselves to $\Re(s) > -2$ and taking the principal branch of log, we end up with a holomorphic function. The fact that $Z_3(s)$ is holomorphic for $\Re(s) > -1$ follows from Hypothesis 3.3 and from the Euler product of the Riemann zeta function (again taking the principal branch of log).

Now Mellin inversion gives that

$$\eta\left(\frac{n}{M}\right) = \frac{1}{2\pi i}\int_{\Re(s)=1} \frac{M^s}{n^s}\mathcal{M}\eta(s)\, ds.$$

Multiplying by $\mathbf{g}_a(n)$ and summing over n yields the result. □

5.1.1. *Properties of the Dirichlet series*

Lemma 5.4. *Assume Hypothesis* 3.3. *We have*

$$Z_3(s) = Z_3(-1) + O(|s+1|)$$

in the region $|s+1| \leqslant 3$, *with* $\Re(s) > -1$. *Note that, by Hypothesis* 3.3, *the product defining* $Z_3(-1)$ *is convergent* (*see the proof of Proposition* 5.9).

Proof. We will show that

$$\log \frac{Z_3(s)}{Z_3(-1)} \ll |s+1|,$$

from which the lemma clearly follows. Let s be a complex number with $\Re(s) > -1$. We compute

$$\begin{aligned}
\log \frac{Z_3(s)}{Z_3(-1)} &= \sum_{p\nmid a} \log\left(\frac{1 + \frac{1}{p^{s+1}}\left(\frac{1}{\gamma(p)} - 1\right)}{\frac{1}{\gamma(p)}} \cdot \frac{\left(1 - \frac{1}{p^{s+2}}\right)^{1-\mathbf{k}}}{\left(1 - \frac{1}{p}\right)^{1-\mathbf{k}}}\right) \\
&= \sum_{p\nmid a}\Bigg[\log\left(1 - (1-\gamma(p))\left(1 - \frac{1}{p^{s+1}}\right)\right) \\
&\quad + (1-\mathbf{k})\log\left(1 + \frac{1}{p-1}\left(1 - \frac{1}{p^{s+1}}\right)\right)\Bigg] \\
&= \sum_{p\nmid a}\left[\frac{\mathbf{h}(p)-1}{p-\mathbf{h}(p)}\left(1 - \frac{1}{p^{s+1}}\right) + \frac{1-\mathbf{k}}{p-1}\left(1 - \frac{1}{p^{s+1}}\right)\right] \\
&\quad + O_\epsilon\left(|s+1|^2 \sum_p \frac{\log^2 p}{p^{2-\epsilon}}\right)
\end{aligned}$$

$$= \sum_{p \nmid a} \left(\frac{\mathbf{h}(p) - 1}{p - \mathbf{h}(p)} + \frac{1 - \mathbf{k}}{p - 1} \right) \left(1 - \frac{1}{p^{s+1}} \right) + O\left(|s+1|^2 \right)$$

$$= \sum_{p} \left(\frac{\mathbf{h}(p) - 1}{p - \mathbf{h}(p)} + \frac{1 - \mathbf{k}}{p - 1} \right) \left(1 - \frac{1}{p^{s+1}} \right) + O\left(|s+1| \right). \tag{13}$$

Note that, by Hypothesis 3.3, the series

$$\sum_{p} \left(\frac{\mathbf{h}(p) - 1}{p - \mathbf{h}(p)} + \frac{1 - \mathbf{k}}{p - 1} \right) = \sum_{p} \frac{\mathbf{h}(p) - \mathbf{k}}{p} + O(1)$$

converges. Moreover, summation by parts yields the following estimate:

$$S(t) := \sum_{p \leqslant t} \left(\frac{\mathbf{h}(p) - 1}{p - \mathbf{h}(p)} + \frac{1 - \mathbf{k}}{p - 1} \right) = S(\infty) + O\left(\frac{1}{\log^2 (t+2)} \right).$$

We then obtain that

$$\begin{aligned}
\sum_{p \leqslant T} \left(\frac{\mathbf{h}(p) - 1}{p - \mathbf{h}(p)} + \frac{1 - \mathbf{k}}{p - 1} \right) \left(1 - \frac{1}{p^{s+1}} \right) &= \int_1^T \left(1 - \frac{1}{t^{s+1}} \right) dS(t) \\
&= \left(1 - \frac{1}{t^{s+1}} \right) S(t) \Big|_1^T - (s+1) \int_1^T \frac{S(t)}{t^{s+2}} dt \\
&= \left(1 - \frac{1}{T^{s+1}} \right) \left(S(\infty) + O\left(\frac{1}{\log^2 T} \right) \right) \\
&\quad - (s+1) \int_1^T \frac{S(\infty)}{t^{s+2}} dt \\
&\quad + O\left(|s+1| \int_1^T \frac{dt}{t \log^2 (t+2)} \right) \\
&= S(\infty) \left(1 - \frac{1}{T^{s+1}} \right) + O\left(\frac{1}{\log^2 T} \right) + \frac{S(\infty)}{t^{s+1}} \Big|_1^T \\
&\quad + O\left(|s+1| \right) \\
&= O\left(\frac{1}{\log^2 T} + |s+1| \right).
\end{aligned}$$

Taking $T \to \infty$ yields that (13) is $\ll |s+1|$. □

LEMMA 5.5. *Let $f(s)$ be a holomorphic function over a domain $\mathcal{D}$. We have that $(f^{(n)}/f)(s)$ is a polynomial in the variables $(f'(s)/f(s))^{(0)}, (f'(s)/f(s))^{(1)}, \ldots, (f'(s)/f(s))^{(n-1)}$, with integer coefficients.*

Proof. The proof goes by induction, using the identity

$$\frac{f^{(n)}}{f} = \left(\frac{f^{(n-1)}}{f} \right)' + \frac{f^{(n-1)}}{f} \frac{f'}{f}.$$ □

LEMMA 5.6. *Assume Hypothesis 3.3. Let $Z_3(s)$ be defined as in (12) and let $n \geqslant 0$. Then there exists $\delta > 0$ such that, uniformly in the region $-1 < \sigma < -\frac{1}{2}$ and $t \in \mathbb{R}$, we have*

$$Z_3^{(n)}(\sigma + it) \ll_n (|t| + 2)^{1/2 - \delta}. \tag{14}$$

In particular, if t is fixed, then $Z_3^{(n)}(\sigma + it)$ is bounded near $\sigma = -1$.

Proof. First write $Z_3(s) = Z_4(s)Z_5(s)$, where

$$Z_4(s) = Z_4(s;a) := \prod_{p \nmid a} \left(1 + \frac{1}{p^{s+1}} \left(\frac{1}{\gamma(p)} - 1\right)\right) \left(1 - \frac{1-\mathbf{k}}{p^{s+2}}\right),$$

$$Z_5(s) = Z_5(s;a) := \prod_{p \nmid a} \left(1 - \frac{1}{p^{s+2}}\right)^{1-\mathbf{k}} \left(1 - \frac{1-\mathbf{k}}{p^{s+2}}\right)^{-1}.$$

The function $Z_5(s)$ is uniformly bounded in the region $\Re(s) \geqslant -1$, since the Eulerian product converges absolutely. As for $Z_4(s)$, we have, for $-1 < \sigma < -\frac{1}{2}$, that

$$\log Z_4(\sigma + it) = \log \prod_{p \nmid a} \left(1 + \frac{1}{p^{\sigma+1+it}} \frac{\mathbf{k} - \mathbf{h}(p)}{p}\right) + O(1).$$

Hypothesis 3.3 gives

$$S(x,t) := \sum_{p \leqslant x} \frac{\mathbf{k} - \mathbf{h}(p)}{p^{1+it}} \leqslant (1/2 - \delta) \log(|t| + 2) + O(1).$$

Thus,

$$\begin{aligned}
\log \prod_{p \nmid a} \left(1 + \frac{1}{p^{\sigma+1+it}} \frac{\mathbf{k} - \mathbf{h}(p)}{p}\right) &= \sum_{p \nmid a} \frac{1}{p^{\sigma+1}} \frac{\mathbf{k} - \mathbf{h}(p)}{p^{1+it}} + O(1) \\
&= \int_1^\infty \frac{dS(x,t)}{x^{\sigma+1}} + O(1) \\
&= \frac{S(x,t)}{x^{\sigma+1}}\Big|_1^\infty + (\sigma+1) \int_1^\infty \frac{S(x,t)}{x^{\sigma+2}} dx + O(1) \\
&\leqslant \left(\frac{1}{2} - \delta\right) \log(|t|+2) \int_1^\infty \frac{\sigma+1}{x^{\sigma+2}} dx + O(1) \\
&= \left(\frac{1}{2} - \delta\right) \log(|t|+2) + O(1),
\end{aligned}$$

which proves (14) for $n = 0$. The bound

$$\sum_{p \leqslant x} \frac{(\mathbf{k} - \mathbf{h}(p)) \log^m p}{p^{1+it}} \ll_\epsilon (|t|+2)^\epsilon$$

gives

$$\left(\frac{Z_3'(\sigma+it)}{Z_3(\sigma+it)}\right)^{(m)} \ll_\epsilon (|t|+2)^\epsilon \tag{15}$$

for $m \geqslant 0$. We complete the proof of (14) for $n \geqslant 1$ by applying Lemma 5.5. □

LEMMA 5.7. *We have, for $|\sigma + it - 1| > \frac{1}{10}$, that*

$$\zeta(\sigma+it) \ll_\epsilon (|t|+2)^{\mu(\sigma)+\epsilon},$$

where

$$\mu(\sigma) = \begin{cases} \frac{1}{2} - \sigma & \text{if } \sigma \leqslant 0, \\ \frac{1}{2} - 2\sigma/3 & \text{if } 0 \leqslant \sigma \leqslant \frac{1}{2}, \\ \frac{1}{3} - \sigma/3 & \text{if } \frac{1}{2} \leqslant \sigma \leqslant 1, \\ 0 & \text{if } \sigma \geqslant 1. \end{cases}$$

Moreover, these bounds are uniform for σ contained in any compact subset of $\mathbb{R}$.

Proof. See [19, Section II.3.4], in particular (II.3.13) and Theorem 3.8. By studying the proof of the Phragment–Lindelöf principle (see [8, Chapter 9] for instance), we see that the bounds we get are uniform in σ. □

LEMMA 5.8. *Assume Hypothesis 3.3. Let*

$$Z(s) = Z(s;a) := \frac{\mathfrak{S}_2(s)\zeta(s+1)\zeta(s+2)^{1-\mathbf{k}}Z_3(s)}{s(s+1)},$$

with $\mathfrak{S}_2(s)$ and $Z_3(s)$ defined as in Lemma 5.3. There exists $\delta > 0$ such that, uniformly for $|t| \geqslant 2$ and $-1 < \sigma < -\frac{1}{2}$,

$$Z^{(n)}(\sigma + it) \ll_n \frac{1}{|t|^{1+\delta}}.$$

Proof. Define

$$Z_6(s) = Z_6(s;a) := \frac{\mathfrak{S}_2(s)\zeta(s+2)^{1-\mathbf{k}}Z_3(s)}{s(s+1)}.$$

Write $s = \sigma + it$, with $-1 < \sigma < -\frac{1}{2}$ and $|t| \geqslant 2$. We have, for $m \geqslant 0$, that

$$\left(\frac{Z_6'(s)}{Z_6(s)}\right)^{(m)} = \left(\frac{\mathfrak{S}_2'(s)}{\mathfrak{S}_2(s)}\right)^{(m)} + (1-\mathbf{k})\left(\frac{\zeta'(s+2)}{\zeta(s+2)}\right)^{(m)} + \left(\frac{Z_3'(s)}{Z_3(s)}\right)^{(m)} - \left(\frac{2s+1}{s(s+1)}\right)^{(m)}.$$

We compute that

$$\left(\frac{\mathfrak{S}_2'(s)}{\mathfrak{S}_2(s)}\right)^{(m)} \ll_m 1, \quad \left(\frac{2s+1}{s(s+1)}\right)^{(m)} \ll_m 1, \quad \left(\frac{Z_3'(s)}{Z_3(s)}\right)^{(m)} \ll_{m,\epsilon} |t|^{\epsilon}.$$

(The first bound is clear, the second follows from the fact that $|t| \geqslant 2$ and the third comes from (15).) Applying Cauchy's formula for the derivatives as in [19, Corollaire II.3.10] and then using the bound [19, (II.3.55)] yields

$$\left(\frac{\zeta'(s+2)}{\zeta(s+2)}\right)^{(m)} \ll_m \log^{m+1}(|t|).$$

Using Lemma 5.5,

$$Z_6^{(m)}(s) \ll_{\epsilon,m} |Z_6(s)||t|^{\epsilon}$$

for $m \geqslant 0$. We now use Lemma 5.6 to bound $|Z_3(s)|$, which gives

$$Z_6^{(m)}(s) \ll_m |\zeta(s+2)^{1-\mathbf{k}}||t|^{-3/2-2\delta}$$

for some $\delta > 0$. Now if $\mathbf{k} \leqslant 1$, we use Lemma 5.7 to bound $\zeta(s+2)^{1-\mathbf{k}}$. Otherwise, we use the bound $(\zeta(s+2))^{-1} \ll \log(|t|)$ (see [19, (II.3.56)]). In both cases, we obtain

$$Z_6^{(m)}(s) \ll_m |t|^{-3/2-\delta}.$$

We now use the Cauchy's formula for the derivatives, which states that

$$\zeta^{(k)}(s+1) = \frac{k!}{2\pi i}\oint_{|z|=r} \zeta(s+1+z)\frac{dz}{z^{k+1}}.$$

Selecting $r = \epsilon/2$ and applying Lemma 5.7, we obtain the bound[†])

$$\zeta^{(k)}(s+1) \ll_{k,\epsilon} |t|^{1/2+\epsilon}.$$

[†]This bound is still valid outside the zero-free region of $\zeta(s+1)$; this is why we considered the ordinary derivatives of $\zeta(s+1)$ instead of its logarithmic derivatives as with the other terms.

We conclude the existence of $\delta > 0$ such that

$$Z^{(n)}(s) = \sum_{i=0}^{n} \binom{n}{i} \zeta^{(i)}(s+1) Z_6^{(n-i)}(s) \ll_n \frac{1}{|t|^{1+\delta}}. \qquad \square$$

5.1.2. *The value of $\mu_{\mathbf{k}}(a, M)$*

Proposition 5.9. *Assume Hypothesis* 3.3. *If $\mathbf{k} \in \mathbb{Z}$, then*

$$\begin{aligned}\frac{1}{2\pi i} \int_{\Re(s)=-\frac{1}{2}} & \frac{\mathfrak{S}_2(s)\zeta(s+1)\zeta(s+2)^{1-\mathbf{k}} Z_3(s)}{s(s+1)} M^s \, ds \\ &= -\frac{\mu_{\mathbf{k}}(a, M)}{M} \left(1 + O\left(\frac{\log\log M}{\log M}\right)\right) \\ &\quad + O_A\left(\frac{1}{M \log^A M}\right)\end{aligned}$$

where $\mu_{\mathbf{k}}(a, M)$ is defined in Definition 3.6.

Proof. We first need to understand the behaviour of

$$\begin{aligned} Z(s) = Z(s; a) &:= \frac{\mathfrak{S}_2(s)\zeta(s+1)\zeta(s+2)^{1-\mathbf{k}} Z_3(s)}{s(s+1)} && (16) \\ &= (s+1)^{\mathbf{k}+\omega_{\mathbf{h}}(a)-2} \frac{\mathfrak{S}_2(s)}{(s+1)^{\omega_{\mathbf{h}}(a)}} \zeta(s+1)((s+1)\zeta(s+2))^{1-\mathbf{k}} \frac{Z_3(s)}{s} && (17)\end{aligned}$$

in the region $\mathcal{D} : -1 \leqslant \Re(s) \leqslant -\frac{1}{2}$. This function is holomorphic for $\Re(s) > -1$ by Lemma 5.3, and, as we will see, the only point in $\mathcal{D}$ where $Z(s)$ is not necessarily locally bounded is $s = -1$. The functions

$$\zeta(s+1), \quad ((s+1)\zeta(s+2))^{1-\mathbf{k}} \quad \text{and} \quad \frac{1}{s}$$

are holomorphic on $\mathcal{D}$ and do not vanish at $s = -1$. The function $Z_3(s)$ is holomorphic for $\Re(s) > -1$, and all its derivatives are locally bounded around any point of $\mathcal{D}$ by Lemma 5.6. We compute

$$Z_3(-1) = \prod_{p \nmid a} \frac{1 - \mathbf{h}(p)/p}{(1 - 1/p)^{\mathbf{k}}} \neq 0,$$

since $\mathbf{h}(p) < p$. As for the function $\mathfrak{S}_2(s)$, it is holomorphic on $\mathcal{D}$. However, this function can vanish at $s = -1$ if, for a certain $p \mid a$, we have $\mathbf{h}(p^f) = \mathbf{h}(p^{f+1})/p$. In this case, we have, for s close to -1, that

$$\begin{aligned}\mathfrak{S}_2(s) \prod_{p \mid a} & \left(1 - \frac{1}{p^{s+2}}\right)^{\mathbf{k}-1} \\ &= \prod_{\substack{p^f \| a: \\ \mathbf{h}(p^f) \neq \mathbf{h}(p^{f+1})/p, \\ f \geqslant 1}} \left[\frac{\mathbf{h}(p^f) - \mathbf{h}(p^{f+1})/p}{1 - 1/p} + O(|s+1|)\right] \\ &\quad \times \prod_{\substack{p^f \| a: \\ \mathbf{h}(p^f) = \mathbf{h}(p^{f+1})/p, \\ f \geqslant 1}} [(s+1)(1 + \mathbf{h}(p) + \cdots + \mathbf{h}(p^f)) \log p + O(|s+1|^2)],\end{aligned}$$

and, since $\mathbf{h}(p^e) \geqslant 0$, this shows that every local factor has at most a simple zero at $s = -1$. We conclude that

$$\frac{\mathfrak{S}_2(s)}{(s+1)^{\omega_{\mathbf{h}}(a)}}$$

is holomorphic on $\mathcal{D}$ and does not vanish at $s = -1$. We now split into three distinct cases, depending on the analytic nature of $(s+1)^{\mathbf{k}+\omega_{\mathbf{h}}(a)-2}$ near $s = -1$.

First case: $\mathbf{k} + \omega_{\mathbf{h}}(a) \geqslant 2$. In this case, $Z(s)$ and all of its derivatives are bounded near $s = -1$. To show this, note that it is true for the functions

$$(s+1)^{\mathbf{k}+\omega_{\mathbf{h}}(a)-2}, \quad \frac{\mathfrak{S}_2(s)}{(s+1)^{\omega_{\mathbf{h}}(a)}}, \quad \zeta(s+1), \quad ((s+1)\zeta(s+2))^{1-\mathbf{k}}, \quad \frac{1}{s} \quad \text{and} \quad Z_3(s),$$

so it is also true for $Z(s)$ by Leibniz's rule. We now shift the contour of integration to the left until the line $\Re(s) = -1 + 1/\log M$ to obtain

$$\begin{aligned}\frac{1}{2\pi i}\int_{\Re(s)=-\frac{1}{2}} Z(s)\,ds &= \frac{i}{2\pi i}\int_{\mathbb{R}} Z\left(-1+\frac{1}{\log M}+it\right) M^{-1+\frac{1}{\log M}+it}\,dt \\ &= \frac{e}{M}\frac{1}{2\pi}\int_{\mathbb{R}} Z\left(-1+\frac{1}{\log M}+it\right) e^{it\log M}\,dt,\end{aligned}$$

which gives, after A integrations by parts,

$$\begin{aligned}\frac{1}{2\pi i}\int_{\Re(s)=-\frac{1}{2}} Z(s)\,ds &\ll_A \frac{1}{M\log^A M}\int_{\mathbb{R}} \left|Z^{(A)}\left(-1+\frac{1}{\log M}+it\right)\right| |e^{it\log M}|\,dt \\ &\ll_A \frac{1}{M\log^A M}\left(O(1)+\int_{|t|\geqslant 2}\frac{1}{|t|^{1+\delta}}dt\right) \\ &\ll_A \frac{1}{M\log^A M}\end{aligned}$$

by Lemma 5.8. Note that the uniformity in σ was crucial. This shows that we can take $\mu_{\mathbf{k}}(a, M) = 0$.

Second case: $\mathbf{k} + \omega_{\mathbf{h}}(a) = 1$. Let

$$c := \lim_{s\to 1^+} (s+1)Z(s) \neq 0$$

and define

$$Z_7(s) = Z_7(s; a) := Z(s) - \frac{c}{s+1}.$$

We can show using Lemma 5.4 that, for s close to -1 with $\Re(s) > -1$, the following bound holds:

$$Z_7(s) \ll 1.$$

Lemma 5.4 implies that, for s close to -1 with $\Re s > -1$, the function

$$Z_7'(s) = \frac{((s+1)Z(s))'}{s+1} - \frac{(s+1)Z(s)}{(s+1)^2} + \frac{c}{(s+1)^2}$$

satisfies

$$Z_7'(s) \ll \frac{1}{|s+1|}.$$

Using Lemma 5.8, we obtain that for $|t| \geqslant 2$,

$$Z_7'(s) \ll \frac{1}{|t|^{1+\delta}}.$$

Thus,

$$\begin{aligned}\frac{1}{2\pi i}\int_{\Re(s)=-1+1/\log M} Z_7(s)M^s\,ds &= \frac{-1}{2\pi i\log M}\int_{\Re(s)=-1+1/\log M} Z_7'(s)M^s\,ds\\ &\ll \frac{1}{M\log M}\left|\int_{-\infty}^{\infty} Z_7'\left(-1+\frac{1}{\log M}+it\right)M^{it}dt\right|\\ &\ll \frac{1}{M\log M}\left(\left|\int_{-2}^{2} Z_7'\left(-1+\frac{1}{\log M}+it\right)M^{it}dt\right|+O(1)\right)\\ &\ll \frac{1}{M\log M}\left(\int_{-2}^{2}\frac{1}{\frac{1}{\log M}+|t|}dt+O(1)\right)\\ &\ll \frac{1}{M\log M}\left(\int_0^{\frac{1}{\log M}}\log M+\int_{\frac{1}{\log M}}^{2}\frac{1}{t}dt+O(1)\right)\\ &\ll \frac{\log\log M}{M\log M}.\end{aligned}$$

Combining this bound with an easy residue computation yields

$$\begin{aligned}\frac{1}{2\pi i}\int_{\Re(s)=-\frac12} Z(s)M^s\,ds &= \frac{1}{2\pi i}\int_{\Re(s)=-\frac12} Z_7(s)M^s\,ds+\frac{1}{2\pi i}\int_{\Re(s)=-\frac12}\frac{c}{s+1}M^s\,ds\\ &= \frac{c}{M}\left(1+O\left(\frac{\log\log M}{\log M}\right)\right).\end{aligned}$$

Now Remarks 4.4 and 4.5 show that $c=-\mu_{\mathbf{k}}(a,M)$, which concludes this case.

Third case: $\mathbf{k}=\omega_{\mathbf{h}}(a)=0$. Defining

$$c:=\lim_{s\to-1^+}(s+1)^2Z(s)\neq 0,$$

we obtain that the function $Z_8(s)=Z_8(s;a):=Z(s)-\frac{c}{(s+1)^2}$ satisfies the bound

$$Z_8(s)\ll\frac{1}{|s+1|}$$

by Lemma 5.4. An easy residue computation yields

$$\frac{1}{2\pi i}\int_{\Re(s)=-\frac12} Z(s)M^s\,ds = c\frac{\log M}{M}+\frac{1}{2\pi i}\int_{\Re(s)=-1+\frac{1}{\log M}} Z_8(s)M^s\,ds.$$

Proceeding in an analogous way to the previous case, we compute

$$\begin{aligned}\int_{\Re(s)=-1+\frac{1}{\log M}} Z_8(s)M^s\,ds &\ll \frac{1}{M}\left|\int_{-\infty}^{\infty} Z_8\left(-1+\frac{1}{\log M}+it\right)M^{it}\,dt\right|\\ &\ll \frac{1}{M}\left(\left|\int_{-2}^{2} Z_8\left(-1+\frac{1}{\log M}+it\right)M^{it}\,dt\right|+O(1)\right)\\ &\ll \frac{\log\log M}{M},\end{aligned}$$

from which we conclude

$$\begin{aligned}\frac{1}{2\pi i}\int_{\Re(s)=-\frac12} Z(s)M^s\,ds &= c\frac{\log M}{M}\left(1+\frac{\log\log M}{\log M}\right)\\ &= -\frac{\mu_0(a,M)}{M}\left(1+\frac{\log\log M}{\log M}\right)\end{aligned}$$

by Remark 4.4, since

$$c = \frac{1}{2} \prod_{\substack{p^f \| a \\ f \geqslant 0}} (\mathbf{h}(p^f) - \mathbf{h}(p^{f+1})/p). \qquad \square$$

LEMMA 5.10. *Let $z > 1$ be a real number. Then*

$$\frac{1}{2\pi i} \int_{\Re(s)=-\frac{1}{2}} \frac{M^s}{(s+1)^z}\, ds = \frac{1}{M} \frac{(\log M)^{z-1}}{\Gamma(z)}.$$

Proof. Let $R \geqslant 2$ be a large real number and consider $\mathcal{H}_R$ a Hankel contour centered at $s = -1$ and truncated at $-R \pm \epsilon i$ (see [19, Théorème II.0.17]). Define C_R to be the union of two circle segments starting at the endpoints of $\mathcal{H}_R$ and ending at the points $\pm iR$. By Cauchy's formula,

$$\begin{aligned}
\frac{1}{2\pi i} \int_{\Re(s)=-\frac{1}{2}} \frac{M^s}{(s+1)^z}\, ds &= \frac{1}{2\pi i} \int_{\Re(s)=0} \frac{M^s}{(s+1)^z}\, ds \\
&= \frac{1}{2\pi i} \int_{\mathcal{H}_R} \frac{M^s}{(s+1)^z}\, ds + \frac{1}{2\pi i} \int_{C_R} \frac{M^s}{(s+1)^z}\, ds \\
&= \frac{1}{2\pi i} \int_{\mathcal{H}_R} \frac{M^s}{(s+1)^z}\, ds + O\left(\frac{1}{R^{z-1}}\right),
\end{aligned}$$

and so, by taking $R \to \infty$,

$$\begin{aligned}
\frac{1}{2\pi i} \int_{\Re(s)=-\frac{1}{2}} \frac{M^s}{(s+1)^z}\, ds &= \frac{1}{2\pi i} \int_{\mathcal{H}_\infty} \frac{M^s}{(s+1)^z}\, ds \\
&= \frac{1}{M} \frac{1}{2\pi i} \int_{\mathcal{H}_\infty} \frac{e^{(s+1)\log M}}{(s+1)^z}\, ds \\
&= \frac{(\log M)^{z-1}}{M} \frac{1}{2\pi i} \int_{\mathcal{H}'_\infty} \frac{e^w}{w^z} dw \\
&= \frac{1}{M} \frac{(\log M)^{z-1}}{\Gamma(z)}
\end{aligned}$$

by Hankel's formula. Here, $\mathcal{H}'_\infty$ denotes an infinite Hankel contour centered at $w = 0$. $\square$

PROPOSITION 5.11. *Assume Hypothesis* 3.3. *If $\mathbf{k} \notin \mathbb{Z}$, then*

$$\frac{1}{2\pi i} \int_{\Re(s)=-\frac{1}{2}} \frac{\mathfrak{S}_2(s)\zeta(s+1)\zeta(s+2)^{1-\mathbf{k}} Z_3(s)}{s(s+1)} M^s\, ds = -\frac{\mu_{\mathbf{k}}(a,M)}{M} \left(1 + O\left(\frac{1}{\log M}\right)\right).$$

Proof. As in Proposition 5.9, we need to study the function

$$Z(s) = Z(s;a) := (s+1)^{\mathbf{k}+\omega_{\mathbf{h}}(a)-2} \frac{\mathfrak{S}_2(s)}{(s+1)^{\omega_{\mathbf{h}}(a)}} \zeta(s+1)((s+1)\zeta(s+2))^{1-\mathbf{k}} \frac{Z_3(s)}{s}$$

in the region $\mathcal{D} : -1 \leqslant \Re(s) \leqslant -\frac{1}{2}$. This function is holomorphic for $\Re(s) > -1$ by Lemma 5.6, and the only point in $\mathcal{D}$ where $Z(s)$ is not necessarily locally bounded is $s = -1$. However, the functions

$$\frac{\mathfrak{S}_2(s)}{(s+1)^{\omega_{\mathbf{h}}(a)}}, \quad \zeta(s+1), \quad ((s+1)\zeta(s+2))^{1-\mathbf{k}} \quad \text{and} \quad \frac{1}{s}$$

are holomorphic on $\mathcal{D}$ and do not vanish at $s = -1$. The function $Z_3(s)$ is holomorphic for $\Re(s) > -1$, all its derivatives are locally bounded around any point of $\mathcal{D}$, and $Z_3(-1) \neq 0$.

Define

$$Z_9(s) = Z_9(s;a) := Z(s) - c(s+1)^{\mathbf{k}+\omega_{\mathbf{h}}(a)-2},$$

where

$$c := \lim_{s\to -1^+} (s+1)^{2-\mathbf{k}-\omega_{\mathbf{h}}(a)} Z(s) \neq 0.$$

We have that

$$\begin{aligned}\frac{1}{2\pi i}\int_{\Re(s)=-\frac12} Z(s)M^s\,ds &= \frac{(-1)^{\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a)}}{2\pi i(\log M)^{\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a)}}\int_{\Re(s)=-\frac12} Z^{(\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a))}(s)M^s\,ds\\
&= \frac{(-1)^{\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a)}}{2\pi i(\log M)^{\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a)}}\Bigg(\int_{\Re(s)=-\frac12} Z_9^{(\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a))}(s)M^s\,ds\\
&\quad + c\frac{\Gamma(\mathbf{k}+\omega_{\mathbf{h}}(a)-1)}{\Gamma(\mathbf{k}-\lceil \mathbf{k}\rceil-1)}\int_{\Re(s)=-\frac12}(s+1)^{\mathbf{k}-\lceil \mathbf{k}\rceil-2}M^s\,ds\Bigg)\\
&= \frac{c}{M}\frac{(\log M)^{1-\mathbf{k}-\omega_{\mathbf{h}}(a)}}{\Gamma(2-\mathbf{k}-\omega_{\mathbf{h}}(a))}\\
&\quad + \frac{(-1)^{\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a)}}{2\pi i(\log M)^{\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a)}}\int_{\Re(s)=-\frac12} Z_9^{(\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a))}(s)M^s\,ds \qquad (18)\end{aligned}$$

by Lemma 5.10. We will show the bound

$$Z_9^{(\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a))}(s) \ll |s+1|^{\mathbf{k}-\lceil \mathbf{k}\rceil-1} \qquad (19)$$

for s close to -1, which will yield (using Lemma 5.8)

$$\begin{aligned}\int_{\Re(s)=-1+1/\log M} Z_9^{(\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a))}(s)M^s\,ds &\ll \frac{1}{M}\left|\int_{-\infty}^{\infty} Z_9^{(\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a))}\left(-1+\frac{1}{\log M}+it\right)M^{it}dt\right|\\
&= \frac{1}{M}\left|\int_{-2}^{2} Z_9^{(\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a))}\left(-1+\frac{1}{\log M}+it\right)M^{it}dt + O(1)\right|\\
&\ll \frac{1}{M}\left(\int_{-2}^{2}\left(\frac{1}{\log M}+|t|\right)^{\mathbf{k}-\lceil \mathbf{k}\rceil-1}dt + O(1)\right)\\
&\ll \frac{1}{M}\left(\int_0^{1/\log M}(\log M)^{1-\mathbf{k}+\lceil \mathbf{k}\rceil} + \int_{1/\log M}^{2} t^{\mathbf{k}-\lceil \mathbf{k}\rceil-1}dt + O(1)\right)\\
&\ll \frac{(\log M)^{\lceil \mathbf{k}\rceil-\mathbf{k}}+1}{M} \ll \frac{(\log M)^{\lceil \mathbf{k}\rceil-\mathbf{k}}}{M},\end{aligned}$$

from which we will conclude, using (18), that

$$\begin{aligned}\frac{1}{2\pi i}\int_{\Re(s)=-\frac12} Z(s)M^s\,ds &= \frac{c}{M}\frac{(\log M)^{1-\mathbf{k}-\omega_{\mathbf{h}}(a)}}{\Gamma(2-\mathbf{k}-\omega_{\mathbf{h}}(a))}\left(1+O\left(\frac{1}{\log M}\right)\right)\\
&= -\mu_{\mathbf{k}}(a,M)\left(1+O\left(\frac{1}{\log M}\right)\right),\end{aligned}$$

achieving the proof. Let us now show that (19) holds. By Lemma 5.6, the function

$$Z_{10}(s) = Z_{10}(s;a) := (s+1)^{2-\mathbf{k}-\omega_{\mathbf{h}}(a)}Z(s)$$

as well as its derivatives are locally bounded around $s=-1$. Moreover, applying Lemma 5.4 gives the bound

$$Z_{10}(s) = Z_{10}(-1) + O(|s+1|). \qquad (20)$$

Now we use Leibniz's formula:

$$\begin{aligned}
Z^{(\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a))}(s) &= \left((s+1)^{\mathbf{k}+\omega_{\mathbf{h}}(a)-2} Z_{10}(s)\right)^{(\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a))}\\
&= \sum_{i=0}^{\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a)} \binom{\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a)}{i} \left((s+1)^{\mathbf{k}+\omega_{\mathbf{h}}(a)-2}\right)^{(i)} Z_{10}^{(\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a)-i)}(s)\\
&= \left((s+1)^{\mathbf{k}+\omega_{\mathbf{h}}(a)-2}\right)^{(\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a))} Z_{10}(s) + O(|s+1|^{\mathbf{k}-\lceil \mathbf{k}\rceil-1})\\
&= \left((s+1)^{\mathbf{k}+\omega_{\mathbf{h}}(a)-2}\right)^{(\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a))} Z_{10}(-1) + O(|s+1|^{\mathbf{k}-\lceil \mathbf{k}\rceil-1})
\end{aligned}$$

by (20), so

$$\begin{aligned}
Z_9^{(\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a))}(s) &= Z^{(\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a))}(s) - c\left((s+1)^{\mathbf{k}+\omega_{\mathbf{h}}(a)-2}\right)^{(\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a))}\\
&= (Z_{10}(-1)-c)\left((s+1)^{\mathbf{k}+\omega_{\mathbf{h}}(a)-2}\right)^{(\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a))} + O(|s+1|^{\mathbf{k}-\lceil \mathbf{k}\rceil-1})\\
&= O(|s+1|^{\mathbf{k}-\lceil \mathbf{k}\rceil-1})
\end{aligned}$$

since $c = Z_{10}(-1)$. $\square$

Lemma 5.12. *Assume Hypothesis* 3.3. *Let $y \geqslant 1$ be a real number. Then*

$$\frac{1}{2\pi i}\int_{\Re(s)=-\frac{1}{2}} \mathfrak{S}_2(s)\zeta(s+1)\zeta(s+2)^{1-\mathbf{k}} Z_3(s) y^s \frac{ds}{s} \ll_\epsilon y^{-1+\epsilon}. \tag{21}$$

Proof. Define

$$Z_{\mathcal{A}}(s) = Z_{\mathcal{A}}(s;a) := \mathfrak{S}_2(s)\zeta(s+1)\zeta(s+2)^{1-\mathbf{k}} Z_3(s).$$

The goal is to bound the integral

$$\frac{1}{2\pi i}\int_{\Re(s)=-\frac{1}{2}} Z_{\mathcal{A}}(s) y^s \frac{ds}{s} = \frac{1}{2\pi i}\int_{\Re(s)=-1+\epsilon} Z_{\mathcal{A}}(s) y^s \frac{ds}{s}.$$

We will first show that this integral is $\ll_\epsilon y^{-1/2+\epsilon}$ using complex analysis, and then we will see how to improve this bound to $\ll_\epsilon y^{-1+\epsilon}$ by elementary means. In the region $-1+\epsilon < \sigma$, we have the bound

$$|Z_{\mathcal{A}}(\sigma+it)| \ll_\epsilon |\zeta(\sigma+1+it)| \ll_\epsilon (|t|+2)^{\mu(\sigma+1)+\epsilon},$$

where $\mu(\sigma+1)$ is defined as in Lemma 5.7. Thus, we obtain the bounds

$$\begin{aligned}
&\int_{-1+\epsilon-iT}^{-1+\epsilon+iT} Z_{\mathcal{A}}(s) y^s \frac{ds}{s} \ll_\epsilon \frac{T^{1/2}}{y^{1-\epsilon}},\\
&\int_{-1+\epsilon\pm iT}^{\epsilon\pm iT} Z_{\mathcal{A}}(s) y^s \frac{ds}{s} \ll_\epsilon (Ty)^\epsilon \left(\frac{1}{T^{5/6}y^{1/2}} + \frac{1}{T^{1/2}y} + \frac{1}{T} + \frac{1}{T^{5/6}y^{1/2}}\right).
\end{aligned}$$

The last integral we need to bound is

$$\begin{aligned}
\frac{1}{2\pi i}\int_{\Re(s)=\epsilon, |\Im(s)|>T} Z_{\mathcal{A}}(s) y^s \frac{ds}{s} &= \sum_n \mathbf{g}_a(n) \frac{1}{2\pi i}\int_{\Re(s)=\epsilon, |\Im(s)|>T} \left(\frac{y}{n}\right)^s \frac{ds}{s}\\
&\ll y^\epsilon \sum_n \mathbf{g}_a(n) \frac{1}{n^\epsilon(1+T|\log(y/n)|)}
\end{aligned}$$

by the effective version of Perron's formula (see [**19**, Théoréme II.2.3]). The last sum is

$$\ll \frac{y^\epsilon}{\sqrt{T}} \sum_{n\leqslant y\left(1-\frac{1}{\sqrt{T}}\right)} \mathbf{g}_a(n) + \sum_{y\left(1-\frac{1}{\sqrt{T}}\right)\leqslant n\leqslant y\left(1+\frac{1}{\sqrt{T}}\right)} \mathbf{g}_a(n) + \frac{y^\epsilon}{\sqrt{T}} \sum_{n\geqslant y\left(1+\frac{1}{\sqrt{T}}\right)} \frac{\mathbf{g}_a(n)}{n^\epsilon}$$
$$\ll \frac{y^\epsilon}{\sqrt{T}}\log y + \frac{1}{\sqrt{T}} \ll_\epsilon \frac{y^\epsilon}{\sqrt{T}}\log y$$

by Lemma 5.2. Taking $T = y$ yields that the left-hand side of (21) is $\ll_\epsilon y^{-1/2+\epsilon}$. We now proceed to show that this bound can be improved to $\ll_\epsilon y^{-1+\epsilon}$. The function $Z_{\mathcal{A}}(s)y^s/s$ has a double pole at $s = 0$ with residue equal to $C_1 \log y + C_2$, where C_1 and C_2 are real numbers independent of y. By the residue theorem and Mellin inversion,

$$\begin{aligned}\frac{1}{2\pi i}\int_{\Re(s)=-\frac{1}{2}} Z_{\mathcal{A}}(s)y^s\frac{ds}{s} &= -C_1\log y - C_2 + \frac{1}{2\pi i}\int_{\Re(s)=1} Z_{\mathcal{A}}(s)y^s\frac{ds}{s} \\ &= \sum_{n\leqslant y}\mathbf{g}_a(n) - C_1\log y - C_2.\end{aligned} \tag{22}$$

Let us give an elementary estimate for the sum appearing on the right-hand side of (22). Define

$$\nu(n) := \prod_{p\mid n}\frac{1-\mathbf{h}(p)}{p-1}.$$

Using the convolution identity

$$\frac{1}{\gamma(n)} = \sum_{rs=n}\mu^2(s)\nu(s),$$

we compute

$$\sum_{n\leqslant y}\frac{\mathbf{f}_a(n)}{n\gamma(n)} = \sum_{n\leqslant y}\mathbf{g}_a(n) = \sum_{s\leqslant y}\frac{\mu^2(s)\nu(s)}{s}\sum_{r\leqslant y/s}\frac{\mathbf{f}_a(rs)}{r} \tag{23}$$

$$\begin{aligned}
&= \sum_{s\leqslant y}\frac{\mu^2(s)\nu(s)}{s}\sum_{(a,s)\mid d\mid a}\mathbf{f}_a(d)\sum_{\substack{r\leqslant y/s:\\ (a,rs)=d}}\frac{1}{r}\\
&= \sum_{s\leqslant y}\frac{\mu^2(s)\nu(s)}{s}\sum_{(a,s)\mid d\mid a}\mathbf{f}_a(d)\sum_{\substack{r\leqslant y/s:\\ (d/(d,s))\mid r\\ (a,rs)=d}}\frac{1}{r}\\
&= \sum_{s\leqslant y}\frac{\mu^2(s)\nu(s)}{s}\sum_{(a,s)\mid d\mid a}\mathbf{f}_a(d)\frac{(d,s)}{d}\sum_{\substack{l\leqslant y(d,s)/ds:\\ (a/d,ls/(d,s))=1}}\frac{1}{l}\\
&= \sum_{s\leqslant y}\frac{\mu^2(s)\nu(s)}{s}\sum_{\substack{(a,s)\mid d\mid a:\\ (a/d,s/(d,s))=1}}\mathbf{f}_a(d)\frac{(d,s)}{d}\sum_{\substack{l\leqslant y(d,s)/ds:\\ (l,a/d)=1}}\frac{1}{l}\\
&= \sum_{s\leqslant y}\frac{\mu^2(s)\nu(s)}{s}\sum_{\substack{(a,s)\mid d\mid a:\\ (a/d,s/(d,s))=1}}\mathbf{f}_a(d)\frac{(d,s)}{d}\frac{\phi(a/d)}{a/d}\left(\log\left(\frac{y(d,s)}{ds}\right)+\gamma\right.\\
&\qquad\left.+\sum_{p\mid a/d}\frac{\log p}{p-1}+O\left(\frac{ds}{y(d,s)}\right)\right).
\end{aligned} \tag{24}$$

Using the bound $\nu(n) \ll_\epsilon n^{-1+\epsilon}$, which is deduced from Hypothesis 3.3, we obtain that the error terms sum to $O_{a,\epsilon}(y^{-1+\epsilon})$. Moreover, we can extend the sum over $s \leqslant y$ to all integers, at the cost of the error term $O_{a,\epsilon}(y^{-1+\epsilon})$. Having done this, (24) becomes

$$\sum_{n \leqslant y} \mathbf{g}_a(n) = \tilde{C}_1 \log y + \tilde{C}_2 + O_{a,\epsilon}(y^{-1+\epsilon}), \tag{25}$$

where $\tilde{C}_1$ and $\tilde{C}_2$ are real numbers that do not depend on y. Substituting (25) into (22) and using our previous bound, we obtain

$$(\tilde{C}_1 - C_1)\log y + \tilde{C}_2 - C_2 + O_{a,\epsilon}(y^{-1+\epsilon}) = \frac{1}{2\pi i}\int_{\Re(s)=-\frac{1}{2}} Z_{\mathcal{A}}(s) y^s \frac{ds}{s} = O_{a,\epsilon}(y^{-1/2+\epsilon}),$$

which of course implies that $\tilde{C}_1 = C_1$ and $\tilde{C}_2 = C_2$ since these numbers do not depend on y. We conclude from (22) and (25) that (21) holds. □

5.1.3. *Proof of Proposition 5.1*

Proof of Proposition 5.1. First, we use Lemma 5.3 to write

$$\begin{aligned} S_5 &:= \sum_{1 \leqslant r \leqslant \mathbf{R}} \mathbf{g}_a(r)\left(1 - \frac{r}{\mathbf{R}}\right) - \sum_{1 \leqslant r \leqslant M} \mathbf{g}_a(r)\left(1 - \frac{r}{M}\right) - \sum_{x/R < q \leqslant x/M} \mathbf{g}_a(q) \\ &= \frac{1}{2\pi i}\int_{\Re(s)=1} \mathfrak{S}_2(s)\zeta(s+1)\zeta(s+2)^{1-\mathbf{k}} Z_3(s)\left(\frac{\mathbf{R}^s - M^s}{s+1} + \left(\frac{x}{\mathbf{R}}\right)^s - \left(\frac{x}{M}\right)^s\right)\frac{ds}{s}. \end{aligned}$$

Writing

$$\psi(s) := \frac{\mathbf{R}^s - M^s}{s+1} + \left(\frac{x}{\mathbf{R}}\right)^s - \left(\frac{x}{M}\right)^s,$$

it is trivial that $\psi(0) = 0$. Using Taylor series, we have, for s close to 0, that

$$\psi(s) = (1 + O(s))(s\log(\mathbf{R}/M) + O(s^2)) + s\log(x/\mathbf{R}) - s\log(x/M) + O(s^2),$$

which means that ψ has a double zero at $s = 0$. Thus,

$$\mathfrak{S}_2(s)\zeta(s+1)\zeta(s+2)^{1-\mathbf{k}} Z_3(s)\frac{\psi(s)}{s}$$

is holomorphic at $s = 0$. Using this fact,

$$\begin{aligned} S_5 &= \frac{1}{2\pi i}\int_{\Re(s)=-\frac{1}{2}} \mathfrak{S}_2(s)\zeta(s+1)\zeta(s+2)^{1-\mathbf{k}} Z_3(s)\psi(s)\frac{ds}{s} \\ &= \frac{1}{2\pi i}\int_{\Re(s)=-\frac{1}{2}} \mathfrak{S}_2(s)\zeta(s+1)\zeta(s+2)^{1-\mathbf{k}} Z_3(s)(\mathbf{R}^s - M^s)\frac{ds}{s(s+1)} + O_\epsilon\left(\left(\frac{\mathbf{R}}{x}\right)^{1-\epsilon}\right) \end{aligned}$$

by Lemma 5.12. We conclude using propositions 5.9 and 5.11 that

$$\begin{aligned} S_5 &= \frac{\mu_{\mathbf{k}}(a, M)}{M}\left(1 + O\left(\frac{\log\log M}{\log M}\right)\right) + O_A\left(\frac{1}{M\log^A M}\right) \\ &\quad - \frac{\mu_{\mathbf{k}}(a, \mathbf{R})}{\mathbf{R}}\left(1 + O\left(\frac{\log\log \mathbf{R}}{\log \mathbf{R}}\right)\right) + O_A\left(\frac{1}{\mathbf{R}\log^A \mathbf{R}}\right) + O_\epsilon\left(\left(\frac{\mathbf{R}}{x}\right)^{1-\epsilon}\right) \\ &= \frac{\mu_{\mathbf{k}}(a, M)}{M}\left(1 + O\left(\frac{\log\log M}{\log M}\right)\right) + O_{A,\delta}\left(\frac{1}{M\log^A M}\right), \end{aligned}$$

since $M(x)^{1+\delta} \leqslant \mathbf{L}(x)^{1+\delta} \leqslant \mathbf{R}(x) \leqslant \sqrt{x}$. □

5.2. *Proofs of theorems 4.1 and 4.1**

We first define the following counting function, which will come in handy for the proofs of this section:

$$\mathcal{A}^*(x;q,a) := \sum_{\substack{|a|<n\leqslant x\\ n\equiv a \bmod q}} \mathbf{a}(n). \tag{26}$$

Proof of Theorem 4.1. Let $M(x)\leqslant \mathbf{L}(x)$ and let $\mathbf{R}=\mathbf{R}(x)$ be as in Hypothesis 3.1. We decompose the sum (9) as follows:

$$\begin{aligned}
&\sum_{q\leqslant x/M} (\mathcal{A}(x;q,a)-\mathbf{a}(a)-\mathbf{g}_a(q)\mathcal{A}(x))\\
&\quad=\sum_{q\leqslant x/M} (\mathcal{A}^*(x;q,a)-\mathbf{g}_a(q)\mathcal{A}(x))+O(1)\\
&\quad=\sum_{x/\mathbf{R}<q\leqslant x} \mathcal{A}^*(x;q,a)-\sum_{x/M<q\leqslant x}\mathcal{A}^*(x;q,a)-\mathcal{A}(x)\sum_{x/\mathbf{R}<q\leqslant x/M}\mathbf{g}_a(q)\\
&\qquad+\sum_{q\leqslant x/\mathbf{R}} (\mathcal{A}^*(x;q,a)-\mathbf{g}_a(q)\mathcal{A}(x))+O(1)\\
&\quad=S_1-S_2-S_3+S_4+O(1).
\end{aligned} \tag{27}$$

Hypothesis 3.4 implies the bound

$$S_4\ll \frac{\mathcal{A}(x)}{M(x)^{1+\delta}}.$$

To evaluate the sums S_1 and S_2, we will interchange divisors in a similar way as in Hooley's paper [**12**]. Setting $n=a+qr$, we have, for positive a, that

$$\begin{aligned}
S_2&=\sum_{x/M<q\leqslant x}\sum_{\substack{|a|<n\leqslant x\\ n\equiv a \bmod q}}\mathbf{a}(n)=\sum_{1\leqslant r<M(x-a)/x}\sum_{\substack{a+rx/M<n\leqslant x\\ n\equiv a \bmod r}}\mathbf{a}(n)\\
&=\sum_{1\leqslant r<M(x-a)/x}\left(\mathcal{A}(x;r,a)-\mathcal{A}\left(a+r\frac{x}{M};r,a\right)\right).
\end{aligned} \tag{28}$$

Using Hypothesis 3.1, we see that there exists $\delta>0$ such that

$$\begin{aligned}
S_2&=\sum_{1\leqslant r<M(x-a)/x}\mathbf{g}_a(r)\left(\mathcal{A}(x)-\mathcal{A}\left(a+r\frac{x}{M}\right)\right)+O\left(\frac{\mathcal{A}(x)}{\mathbf{L}(x)^{1+2\delta}}\right)\\
&=\sum_{1\leqslant r<M(x-a)/x}\mathbf{g}_a(r)\left(\mathcal{A}(x)-\mathcal{A}\left(\frac{r}{M}x\right)\right)+O\left(\frac{\mathcal{A}(x)}{\mathbf{L}(x)^{1+\delta}}\right)\\
&=\mathcal{A}(x)\sum_{1\leqslant r<M(x-a)/x}\mathbf{g}_a(r)\left(1-\frac{\mathcal{A}\left(\frac{r}{M}x\right)}{\mathcal{A}(x)}\right)+O\left(\frac{\mathcal{A}(x)}{\mathbf{L}(x)^{1+\delta}}\right)
\end{aligned} \tag{29}$$

by Hypotheses 3.2 and Lemma 5.2. Now, if a were negative, we would have to add an error term of size $\ll \mathcal{A}(x)/\mathbf{L}(x)^{1+\delta}$ to (28) (by Hypothesis 3.2), which would yield the same error term in (29). Using Hypothesis 3.2 again, (29) becomes

$$=\mathcal{A}(x)\sum_{1\leqslant r<M(x-a)/x}\mathbf{g}_a(r)\left(1-\frac{r}{M}\right)+O\left(\frac{\mathcal{A}(x)}{\mathbf{L}(x)^{1+\delta}}\right).$$

If M is an integer, then the Mth term of the sum is $\mathbf{g}_a(r)(1 - M/M) = 0$. If not, the bound $\mathbf{g}_a(r) \ll_\epsilon 1/\phi(r)$ (see Lemma 5.2) implies that this last term is $\ll \mathcal{A}(x)(\log\log M/M^2)$. Thus,

$$S_2 = \mathcal{A}(x) \sum_{1 \leqslant r \leqslant M} \mathbf{g}_a(r) \left(1 - \frac{r}{M}\right) + O\left(\frac{\mathcal{A}(x)}{M^{1+\delta}}\right)$$

since $M(x) \leqslant \mathbf{L}(x)$. A similar calculation shows that

$$S_1 = \mathcal{A}(x) \sum_{1 \leqslant r \leqslant \mathbf{R}(x)} \mathbf{g}_a(r) \left(1 - \frac{r}{\mathbf{R}(x)}\right) + O\left(\frac{\mathcal{A}(x)}{\mathbf{L}(x)^{1+\delta}}\right).$$

Grouping terms, (27) becomes

$$\begin{aligned}
&\sum_{q \leqslant \frac{x}{M}} (\mathcal{A}(x;q,a) - \mathbf{a}(a) - \mathbf{g}_a(q)\mathcal{A}(x)) = S_1 - S_2 - S_3 + S_4 + O(1) \\
&= \mathcal{A}(x) \left(\sum_{1 \leqslant r \leqslant \mathbf{R}} \mathbf{g}_a(r)\left(1 - \frac{r}{\mathbf{R}}\right) - \sum_{1 \leqslant r \leqslant M} \mathbf{g}_a(r)\left(1 - \frac{r}{M}\right) - \sum_{\frac{x}{\mathbf{R}} < q \leqslant \frac{x}{M}} \mathbf{g}_a(q) \right) \\
&\quad + O\left(\frac{\mathcal{A}(x)}{M^{1+\delta}}\right),
\end{aligned}$$

which combined with Proposition 5.1 gives

$$= \frac{\mathcal{A}(x)}{M} \mu_{\mathbf{k}}(a, M) \left(1 + O\left(\frac{\log\log M}{\log M}\right)\right) + O_A\left(\frac{\mathcal{A}(x)}{M \log^A M}\right),$$

that is,

$$\begin{aligned}
&\sum_{q \leqslant x/M} (\mathcal{A}(x;q,a) - \mathbf{a}(a) - \mathbf{g}_a(q)\mathcal{A}(x)) \\
&\quad = \frac{\mathcal{A}(x)}{M} \left(\mu_{\mathbf{k}}(a, M) \left(1 + O\left(\frac{\log\log M}{\log M}\right)\right) + O_A\left(\frac{1}{\log^A M}\right)\right). \qquad \square
\end{aligned}$$

Proof of Theorem 4.1*. Let $M(x) \leqslant \mathbf{L}(x)$ and let $\mathbf{R} = \mathbf{R}(x)$ be as in Hypothesis 3.4. We decompose the sum (10) as follows:

$$\begin{aligned}
&\sum_{x/2M < q \leqslant x/M} (\mathcal{A}(x;q,a) - \mathbf{a}(a) - \mathbf{g}_a(q)\mathcal{A}(x)) \\
&= \sum_{x/2M < q \leqslant x/M} (\mathcal{A}^*(x;q,a) - \mathbf{g}_a(q)\mathcal{A}(x)) + O(1) \\
&= \sum_{x/2M < q \leqslant x} \mathcal{A}^*(x;q,a) - \sum_{x/M < q \leqslant x} \mathcal{A}^*(x;q,a) - \mathcal{A}(x) \sum_{x/2M < q \leqslant x/M} \mathbf{g}_a(q) + O(1) \\
&= S_1 - S_2 - S_3 + O(1).
\end{aligned} \tag{30}$$

Arguing as in the proof of Theorem 4.1, we set $n = a + qr$ to obtain that, for positive a,

$$\begin{aligned}
S_2 &= \sum_{1 \leqslant r < M(x-a)/x} \left(\mathcal{A}(x;r,a) - \mathcal{A}\left(a + r\frac{x}{M}; r, a\right)\right) \\
&= \mathcal{A}(x) \sum_{1 \leqslant r < M(x-a)/x} \mathbf{g}_a(r) \left(1 - \frac{\mathcal{A}((r/M)x)}{\mathcal{A}(x)}\right) + O\left(\frac{\mathcal{A}(x)}{\mathbf{L}(x)^{1+\delta}}\right) \\
&= \mathcal{A}(x) \sum_{1 \leqslant r \leqslant M} \mathbf{g}_a(r) \left(1 - \frac{r}{M}\right) + O\left(\frac{\mathcal{A}(x)}{M^{1+\delta}}\right)
\end{aligned}$$

by Hypotheses 3.1*, 3.2 and Lemma 5.2. Now, if a were negative, we would have to add a negligible contribution. Thus, (30) becomes

$$\begin{aligned}\sum_{x/2M<q\leqslant x/M} &(\mathcal{A}(x;q,a)-\mathbf{a}(a)-\mathbf{g}_a(q)\mathcal{A}(x))\\ &=\mathcal{A}(x)\left(\sum_{1\leqslant r\leqslant 2M}\mathbf{g}_a(r)\left(1-\frac{r}{2M}\right)\right.\\ &\quad\left.-\sum_{1\leqslant r\leqslant M}\mathbf{g}_a(r)\left(1-\frac{r}{M}\right)-\sum_{x/2M<q\leqslant x/M}\mathbf{g}_a(q)\right)+O\left(\frac{\mathcal{A}(x)}{M^{1+\delta}}\right).\end{aligned}$$

Going through the proof of Proposition 5.1, we see that this is

$$\begin{aligned}&=\frac{\mathcal{A}(x)}{M}\mu_{\mathbf{k}}(a,M)\left(1+O\left(\frac{\log\log M}{\log M}\right)\right)-\frac{\mathcal{A}(x)}{2M}\mu_{\mathbf{k}}(a,2M)\left(1+O\left(\frac{\log\log M}{\log M}\right)\right)\\ &\quad+O_A\left(\frac{\mathcal{A}(x)}{M\log^A M}\right),\end{aligned}$$

that is,

$$\begin{aligned}\sum_{x/2M<q\leqslant x/M}&(\mathcal{A}(x;q,a)-\mathbf{a}(a)-\mathbf{g}_a(q)\mathcal{A}(x))\\ &=\frac{\mathcal{A}(x)}{2M}\left(\mu_{\mathbf{k}}(a,M)\left(1+O\left(\frac{\log\log M}{\log M}\right)\right)+O_A\left(\frac{1}{\log^A M}\right)\right),\end{aligned}$$

since, by the definition of $\mu_{\mathbf{k}}(a,M)$,

$$2\mu_{\mathbf{k}}(a,M)-\mu_{\mathbf{k}}(a,2M)=\mu_{\mathbf{k}}(a,M)\left(1+O\left(\frac{1}{\log M}\right)\right).\qquad\square$$

6. *Further Proofs*

In this section, we prove the results of Section 2.

Proof of Theorem 2.2. Put

$$a(n):=\Lambda(n),$$

which gives $\mathcal{A}(x)=\psi(x)$ and $\mathcal{A}(x;q,a)=\psi(x;q,a)$. Define

$$\mathbf{f}_a(q):=\begin{cases}1 & \text{if } (a,q)=1,\\ 0 & \text{otherwise,}\end{cases}$$

and $\gamma(q):=\phi(q)/q$. Define also the multiplicative function $\mathbf{h}(d)$ by $\mathbf{h}(1)=1$, and $\mathbf{h}(d)=0$ for $d>1$. The prime number theorem in arithmetic progressions gives the asymptotic

$$\mathcal{A}(x;q,a)\sim\frac{\mathbf{f}_a(q)}{q\gamma(q)}\mathcal{A}(x),$$

for any fixed a and q such that $(a,q)=1$. Now let us show that the hypotheses of Section 3.2 hold. Fix $A>0$ and put $\mathbf{L}(x):=(\log x)^A$, $\mathbf{R}(x):=x^{1/2}(\log x)^{-B(A)}$, where $B(A):=A+5$. Hypothesis 3.1 is the Bombieri–Vinogradov theorem. Hypothesis 3.2 follows from the prime number theorem. As $\mathbf{h}(p)=\mathbf{k}=0$, Hypothesis 3.3 is trivial. Hypothesis 3.4 follows from [**3**, Theorem 9].

We now compute $\mu_{\mathbf{k}}(a, M)$. As $\mathbf{h}(p^e) = 0$, we have $\omega_{\mathbf{h}}(a) = \omega(a)$, the number of distinct prime factors of a. Thus, Remark 4.4 gives

$$\mu_0(a, M) = \begin{cases} -\dfrac{1}{2}\log M & \text{if } a = \pm 1, \\ -\dfrac{1}{2}\left(1 - \dfrac{1}{p}\right)\log p & \text{if } a = \pm p^e, \\ 0 & \text{if } \omega(a) \geqslant 2, \end{cases}$$

so an application of Theorem 4.1 gives the result with a weaker error term. A better version of Proposition 5.1 follows from Huxley's subconvexity result [**13**], yielding the stated error term (see [**9**] for a more precise proof). □

Proof of Theorem 2.4. Let $Q(x, y) := \alpha x^2 + \beta xy + \gamma y^2$ be a binary quadratic form, where α, β and γ are integers such that $\alpha > 0$, $(\alpha, \beta, \gamma) = 1$ and $d := \beta^2 - 4\alpha\gamma < 0$ (so $Q(x, y)$ is positive definite). Note that the set of d for which $d \equiv 1, 5, 9, 12, 13 \bmod 16$ includes a large subset of all fundamental discriminants. The set of bad primes is $\mathcal{S} := \{p : p \mid 2d\}$ in this case. Since $\mathcal{S} \neq \emptyset$, we will need to modify the proof of Theorem 4.1. We define

$$\chi_d := \left(\frac{4d}{\cdot}\right).$$

Note that, for $(n, 2d) = 1$, we have the equalities

$$r_d(n) = \sum_{m \mid n} \chi_d(m) = \prod_{\substack{p^k \| n: \\ \chi_d(p) = 1}} (k+1) \prod_{\substack{p^k \| n: \\ \chi_d(p) = -1, \\ k \text{ odd}}} 0. \tag{31}$$

An intuitive argument suggests that

$$\mathcal{A}(x; q, a) \sim \frac{R_a(q)}{q^2}\mathcal{A}(x),$$

where

$$R_a(q) := \#\{1 \leqslant x, y \leqslant q : Q(x, y) \equiv a \bmod q\}. \tag{32}$$

As this is a classical result, we leave its proof, as well as several other classical facts about binary quadratic forms, to Appendix A. The function

$$\mathbf{g}_a(q) := \frac{R_a(q)}{q^2}$$

is actually multiplicative (see Lemma A.1), and Lemma A.3 shows that, for $p \nmid 2d$, $\mathbf{g}_a$ is given as in (6) with

$$\mathbf{h}(p^e) := \begin{cases} 1 + e\left(1 - \dfrac{1}{p}\right) & \text{if } \chi_d(p) = 1, \\ \dfrac{1}{p} & \text{if } \chi_d(p) = -1 \text{ and } 2 \nmid e, \\ 1 & \text{if } \chi_d(p) = -1 \text{ and } 2 \mid e, \end{cases}$$

and, for $p \mid 2d$, $R_a(p^e)$ is given as in (A.7) and (A.8). Since we are looking at large moduli, we need to use a result of Plaksin [**18**, Lemma 8], which asserts that

$$\mathcal{A}(x; q, a) = \mathbf{g}_a(q)\mathcal{A}(x) + E(x, q), \tag{33}$$

where $E(x, q) \ll_{a,\epsilon} (x/q)^{3/4+\epsilon}$ if $q \leqslant x^{1/3}$, and $E(x, q) \ll_{a,\epsilon} x^{2/3+\epsilon}q^{-1/2}$ if $x^{1/3} < q \leqslant x^{2/3}$. Summing (33) over $q \leqslant x^{1/2}$, we obtain that the Hypotheses 3.1 and 3.4 hold with $\mathbf{R}(x) := x^{1/2}$ and $\mathbf{L}(x) := x^{\lambda}$, provided $\lambda < \frac{1}{12}$. (Note that, in the case $\beta = 0$, we can take the wider range

$\lambda < \frac{1}{8}$, using [**17**, Lemma 20].) Hypothesis 3.2 follows from Gauss's estimate:

$$\mathcal{A}(x) = A_Q x + O(x^{1/2}),$$

where A_Q is the area of the region $\{(x,y) \in \mathbb{R}^2_{\geqslant 0} : Q(x,y) \leqslant 1\}$. Let us turn to Hypothesis 3.3. For $p \nmid 2d$,

$$\mathbf{h}(p) = \begin{cases} 2 - \dfrac{1}{p} & \text{if } \chi_d(p) = 1, \\ \dfrac{1}{p} & \text{if } \chi_d(p) = -1, \end{cases}$$

so we set $\mathbf{k} := 1$ and

$$\sum_{p \notin \mathcal{S}} \frac{\mathbf{h}(p) - \mathbf{k}}{p} = \sum_{p \nmid 2d} \frac{\chi_d(p)}{p} + O(1) < \infty$$

by the prime number theorem for $\psi(x, \chi_{-d})$ (see [**7**]). Moreover,

$$\begin{aligned} \sum_{p \notin \mathcal{S}} \frac{(\mathbf{h}(p) - \mathbf{k})(\log p)^{n+1}}{p^{1+it}} &= O(1) + (-1)^{n+1} \left(\frac{L'}{L}\right)^{(n)} (1 + it, \chi_d) \\ &\ll_{d,n} (\log(|t| + 2))^{n+2}, \end{aligned} \tag{34}$$

this last bound following from Cauchy's formula for the derivatives combined with the classical bound for $L'(s,\chi)/L(s,\chi)$ in a zero-free region (see [**7**, Chapter 19]). As in the [**19**, proof of Théorème II.3.22], we can deduce from (34) that (setting $\eta := 1/\log^2(|t| + 2)$)

$$\begin{aligned} \sum_{p \notin \mathcal{S}} \frac{\mathbf{h}(p) - \mathbf{k}}{p^{1+it}} + O(1) &= \log L(1 + it, \chi_d) \\ &= \int_{1+it+\eta}^{1+it} \frac{L'(s,\chi_d)}{L(s,\chi_d)} \, ds + \log L(1 + it + \eta, \chi_d) \\ &\ll \eta \log^2(|t| + 2) + \log \zeta(1 + \eta) \\ &= 2 \log\log(|t| + 2) + O(1). \end{aligned}$$

Having proved Hypotheses 3.1–3.4, we now proceed to prove an analogue of Theorem 4.1 (since $\mathcal{S} = \{p : p \mid 2d\}$ is non-empty). In the proof of Lemma 5.3, we need to change the definition of $\mathfrak{S}_2(s)$ to (remember that $(a, 2d) = 1$)

$$\begin{aligned} \mathfrak{S}_2(s) = &\left(\left(1 - \frac{1}{2^{s+1}}\right)\left(1 + \frac{R_a(2)}{2^{s+2}}\right) + \frac{R_a(4)}{4} \frac{1}{2^{2s+2}}\right) \prod_{\substack{p \mid d \\ p \neq 2}} \left(1 - \frac{1}{p^{s+1}} + \frac{R_a(p)}{p^{s+2}}\right) \\ &\times \prod_{\substack{p^f \parallel a \\ f \geqslant 1 \\ p \notin \mathcal{S}}} \left[\left(1 + \frac{\mathbf{h}(p)}{p^{s+1}} + \cdots + \frac{\mathbf{h}(p^f)}{p^{f(s+1)}}\right)\left(1 - \frac{1}{p^{s+1}}\right) + \frac{\mathbf{h}(p^f) - \mathbf{h}(p^{f+1})/p}{1 - 1/p} \frac{1}{p^{(f+1)(s+1)}}\right], \end{aligned}$$

(We also need to change the condition on the product defining $Z_3(s)$ to $p \nmid 2ad$) so

$$\begin{aligned} \mathfrak{S}_2(-1) &= \frac{R_a(4)}{4} \prod_{\substack{p \mid d \\ p \neq 2}} \frac{R_a(p)}{p} \prod_{\substack{p^f \parallel a \\ f \geqslant 1 \\ p \notin \mathcal{S}}} \frac{\mathbf{h}(p^f) - \mathbf{h}(p^{f+1})/p}{1 - 1/p} \\ &= \frac{R_a(4)}{4} \prod_{\substack{p^f \parallel d \\ p \neq 2}} \frac{R_a(p^f)}{p^f} \prod_{\substack{p^f \parallel a: \\ \chi_d(p) = 1}} \left(1 - \frac{1}{p}\right)(f + 1) \prod_{\substack{p^f \parallel a: \\ \chi_d(p) = -1, \\ f \text{ even}}} \left(1 + \frac{1}{p}\right) \prod_{\substack{p^f \parallel a: \\ \chi_d(p) = -1, \\ f \text{ odd}}} 0 \end{aligned}$$

$$= \frac{R_a(4d)}{4d} \prod_{p|a} \left(1 - \frac{\chi_d(p)}{p}\right) r_d(|a|),$$

by (31) and Lemma A.3. We conclude that Theorem 4.1 holds with

$$\mu_1(a, M) = -\frac{R_a(4d)}{4d} \cdot \frac{r_d(|a|)}{2L(1, \chi_d)},$$

which gives the result (with a weaker error term) by Dirichlet's class number formula. To get the better error term $O_\epsilon(1/M^{1/3-\epsilon})$, one has to get a better estimate in Proposition 5.1. To do this, we go back to the proof of Proposition 5.9 and remark that (with the notation introduced there)

$$Z_3(s) = \prod_{p\nmid 2ad} \left(1 - \frac{\chi_d(p)}{p^{s+2}}\right),$$

so

$$Z(s) = \frac{\mathfrak{S}_3(s)\zeta(s+1)L(s+2, \chi_d)}{s(s+1)},$$

where

$$\mathfrak{S}_3(s) := \mathfrak{S}_2(s) \prod_{p|2ad} \left(1 - \frac{\chi_d(p)}{p^{s+2}}\right)^{-1}.$$

Since $Z(s)$ is a meromorphic function on the whole complex plane, we can shift the contour of integration to the left until the line $\Re(s) = -\frac{4}{3} + \epsilon$. We have the following convexity bound on $L(s, \chi_d)$ for $0 \leqslant \sigma \leqslant 1$ and $t \in \mathbb{R}$:

$$L(\sigma + it, \chi_d) \ll_\epsilon (d(|t| + 3))^{(1-\sigma)/2+\epsilon}$$

(see [**15**, (5.20)]). Combining this bound with a standard residue calculation yields that

$$\frac{1}{2\pi i} \int_{\Re(s)=-\frac{1}{2}} \frac{\mathfrak{S}_3(s)\zeta(s+1)L(s+2, \chi_d)}{s(s+1)} M^s \, ds = -\frac{\mu_1(a, M)}{M} + O_\epsilon\left(\frac{1}{M^{4/3-\epsilon}}\right),$$

from which we conclude the result. □

Proof of Theorem 2.6. Set $\mathcal{S} := \{2\}$, $\mathbf{k} := \frac{1}{2}$ and $\mathbf{L}(x) := (\log x)^\lambda$ with $\lambda < \frac{1}{5}$. We first prove Hypothesis 3.2 using a refinement of a theorem of Landau. We have

$$\mathcal{A}(x) = C \frac{x}{\sqrt{\log x}} \left(1 + O\left(\frac{x}{\log x}\right)\right), \tag{35}$$

with

$$C := \frac{1}{\sqrt{2}} \prod_{p \equiv 3 \bmod 4} \left(1 - \frac{1}{p^2}\right)^{-1/2}.$$

(See, for instance, [**19**, Exercice 240].) The distribution of $\mathcal{A}$ in the arithmetic progressions $a \bmod q$ with $(a, q) = 1$ is uniform; however, a result of the strength of Plaksin's (33) is far from being known. The best result so far for individual values of q (in terms of uniformity in q) is due to Iwaniec [**14**], which proved using the half-dimensional sieve that if $(a, q) = 1$ and $a \equiv 1 \bmod (q, 4)$, then

$$\mathcal{A}(x; q, a) = \frac{(2, q)}{(4, q)q\gamma(q)} \mathcal{A}(x) \left(1 + O\left(\frac{\log q}{\log x}\right)^{1/5}\right), \tag{36}$$

where

$$\gamma(q) := \prod_{\substack{p \mid q \\ p \equiv 3 \bmod 4}} \left(1 + \frac{1}{p}\right)^{-1}.$$

An easy computation using the arithmetic properties of $\mathbf{a}(n)$ shows that

$$\mathcal{A}_{p^e}(x) = \begin{cases} \mathcal{A}\left(\dfrac{x}{p^{e+1}}\right) & \text{if } p \equiv 3 \bmod 4 \text{ and } 2 \nmid e, \\ \mathcal{A}\left(\dfrac{x}{p^e}\right) & \text{otherwise,} \end{cases}$$

and more generally,

$$\mathcal{A}_d(x) = \mathcal{A}\left(\frac{\mathbf{h}(d)}{d} x\right), \tag{37}$$

with

$$\mathbf{h}(p^e) := \begin{cases} \dfrac{1}{p} & \text{if } p \equiv 3 \bmod 4 \text{ and } 2 \nmid e, \\ 1 & \text{otherwise.} \end{cases}$$

This confirms that our choice of $\mathbf{k} = \frac{1}{2}$ was good, and Hypothesis 3.3 follows as in the proof of Theorem 2.4. Moreover, (36) can be extended to $(a, q) = d$ for any fixed odd integer $d > 1$, by using the identity $\mathcal{A}(x; q, a) = \mathcal{A}((\mathbf{h}(d)/d)x; q/d, a/d)$, hence Hypothesis 3.1* holds. As we have shown every hypothesis, we turn to the calculation of the average $\mu_{1/2}(a, M)$ (which is never zero since $\mathbf{k} \notin \mathbb{Z}$). We need to modify the definition of $\mathfrak{S}_2(s)$, changing the local factor at $p = 2$ to

$$\left(1 - \frac{1}{2^{s+2}}\right)^{1/2} \left(1 - \frac{1}{2^{2s+2}} + \frac{1}{2^{2s+3}}\right).$$

Doing so and proceeding as in the proof of Theorem 4.1*, we obtain the result. □

LEMMA 6.1. *Suppose that* $\mathcal{H} = \{a_1 n + b_1, \dots a_k n + b_k\}$ *is an admissible* k*-tuple of linear forms and* q, a *are two integers such that* $(q, a_i a + b_i) = 1$ *for* $1 \leqslant i \leqslant k$. *Then the modified* k*-tuple* $\tilde{\mathcal{H}} := \{a_1(qm + a) + b_1, \dots a_k(qm + a) + b_k\}$ *is also admissible. Moreover,*

$$\mathfrak{S}(\tilde{\mathcal{H}}) = \prod_{p \mid q} \left(1 - \frac{\nu_{\mathcal{H}}(p)}{p}\right)^{-1} \mathfrak{S}(\mathcal{H}).$$

Proof. First, since $\mathcal{H}$ is admissible, we have $(a_i, b_i) = 1$ for $1 \leqslant i \leqslant k$. Fix a prime p. We need to show that $\nu_{\tilde{\mathcal{H}}}(p) < p$. For a fixed i we have either $p \mid a_i$, in which case $p \nmid b_i$ so $a_i n + b_i \not\equiv 0 \bmod p$, or $p \nmid a_i$, in which case the only solution to $a_i n + b_i \equiv 0 \bmod p$ is $n \equiv -a_i^{-1} b_i$. Hence, if $p \nmid a_i$, then there are only $\nu_{\mathcal{H}}(p) < p$ distinct possible values for $-a_i^{-1} b_i \bmod p$; thus regrouping these, we can write

$$\prod_{i=1}^{k} (a_i n + b_i) \equiv C \prod_{i : p \mid a_i} b_i \prod_{j=1}^{\nu_{\mathcal{H}}(p)} (n + k_j)^{e_j} \bmod p,$$

where the k_j are distinct integers, $e_j \geqslant 1$ and $p \nmid C$. Using this and the fact that $(a_i, b_i) = 1$, we obtain

$$\prod_{i=1}^{k} (a_i(qm + a) + b_i) \equiv D \prod_{j=1}^{\nu_{\mathcal{H}}(p)} (qm + a + k_j)^{e_j} \bmod p,$$

with $p \nmid D$. If $p \nmid q$, then this has exactly $\nu_{\mathcal{H}}(p) < p$ solutions, therefore $\nu_{\tilde{\mathcal{H}}}(p) < p$. Otherwise, this becomes

$$\prod_{i=1}^{k}(a_i(qm+a)+b_i) \equiv \prod_{i=1}^{k}(a_ia+b_i) \not\equiv 0 \bmod p,$$

since $(q, a_ia+b_i) = 1$ for $1 \leqslant i \leqslant k$. We conclude that $\tilde{\mathcal{H}}$ is admissible. The calculation of $\mathfrak{S}(\tilde{\mathcal{H}})$ follows easily. □

Proof of Theorem 2.8. Define $\mathcal{S} := \emptyset$ and

$$\mathbf{a}(n) := \prod_{\mathcal{L}\in\mathcal{H}} \Lambda(\mathcal{L}(n)) = \Lambda(a_1n+b_1)\Lambda(a_2n+b_2)\cdots\Lambda(a_kn+b_k).$$

In our context, some assumptions of Section 3.1 do not hold. The reason is that the asymptotic for $\mathcal{A}(x;q,a)$ depends on $(q,\mathcal{P}(a;\mathcal{H}))$ rather than depending only on (q,a). The correct conjecture in this case is that for integers a and q such that $(q,\mathcal{P}(a;\mathcal{H})) = 1$, (see [16]†)

$$\mathcal{A}(x;q,a) \sim \frac{\mathcal{A}(x)}{q\gamma(q)},$$

with

$$\gamma(q) := \prod_{p\mid q}\left(1-\frac{\nu_{\mathcal{H}}(p)}{p}\right).$$

This actually follows from the Hardy–Littlewood conjecture, by taking the modified k-tuple of linear forms $\tilde{\mathcal{L}}_i(m) := a_i(qm+a)+b_i = qa_im+aa_i+b_i$, which is admissible if $\mathcal{H}$ is and $(q,\mathcal{P}(a;\mathcal{H})) = 1$ (see Lemma 6.1). Using this idea, we obtain that the assumption of (5) holding uniformly for $|a_i| \leqslant \mathbf{L}(x)^{1+\delta}$ implies Hypothesis 3.1*. We now prove an analogue of Proposition 5.1. Defining

$$Z_{\mathcal{H}}(s) = Z_{\mathcal{H}}(s;a) := \sum_n \frac{\mathbf{f}_a(n)}{n^{s+1}\gamma(n)},$$

where

$$\mathbf{f}_a(q) := \begin{cases} 1 & \text{if } (\mathcal{P}(a;\mathcal{H}),q) = 1, \\ 0 & \text{otherwise,}\end{cases}$$

one can compute that

$$Z_{\mathcal{H}}(s) = \mathfrak{S}_2(s)\zeta(s+1)\zeta(s+2)^kZ_0(s)$$

with

$$\mathfrak{S}_2(s) := \prod_{p\mid\mathcal{P}(a;\mathcal{H})}\left(1-\frac{1}{p^{s+1}}\right)\left(1+\frac{\nu_{\mathcal{H}}(p)}{p-\nu_{\mathcal{H}}(p)}\frac{1}{p^{s+1}}\right)^{-1}$$

and

$$Z_0(s) := \prod_p\left(1+\frac{\nu_{\mathcal{H}}(p)}{p-\nu_{\mathcal{H}}(p)}\frac{1}{p^{s+1}}\right)\left(1-\frac{1}{p^{s+2}}\right)^k,$$

which converges for $\Re(s) > -\frac{3}{2}$. Note that $Z_{\mathcal{H}}(s)$ has a simple pole at $s=0$. Also, $\mathfrak{S}_2(s)$ has a zero of order $\omega(\mathcal{P}(a,\mathcal{H}))$ at the point $s=-1$, and $Z_0(-1) = \mathfrak{S}(\mathcal{H})^{-1}$, so $Z_{\mathcal{H}}(s)$ is of order

†Kawada imposes the additional condition that $R(\mathbf{b}) := \prod_{j=1}^k |a_j| \prod_{1\leqslant i,j\leqslant k}|a_ib_j - a_jb_i|$ is non-zero. However, we assume that our linear forms are admissible and distinct, and $a_i \geqslant 1$; one can show that this implies $R(\mathbf{b}) \neq 0$.

$\omega(\mathcal{P}(a,\mathcal{H}))-k$ at this point. The function

$$\psi(s) := \frac{(2M)^s - M^s}{s+1} + \left(\frac{x}{2M}\right)^s - \left(\frac{x}{M}\right)^s$$

vanishes to the second order at $s=0$. Combining all this information, we obtain, by shifting the contour of integration to the left, that

$$\frac{1}{2\pi i}\int_{\Re(s)=1} Z_{\mathcal{H}}(s)\psi(s)\frac{ds}{s} = \frac{1}{2M}\left(\mu_{1-k}(a,M)(1+o(1)) + O\left(\frac{1}{M^{\delta_k}}\right)\right),$$

where

$$\mu_{1-k}(a,M) := \begin{cases} -\dfrac{1}{2\mathfrak{S}(\mathcal{H})}\dfrac{(\log M)^{k-\omega(\mathcal{P}(a;\mathcal{H}))}}{(k-\omega(\mathcal{P}(a;\mathcal{H})))!}\displaystyle\prod_{p\mid\mathcal{P}(a;\mathcal{H})}\frac{p-\nu_{\mathcal{H}}(p)}{p}\log p & \text{if } \omega(\mathcal{P}(a;\mathcal{H}))\leqslant k,\\ 0 & \text{otherwise,}\end{cases}$$

and $\delta_k>0$ is a small real number (one can take $\delta_k=\frac{1}{2+k}$). We conclude by proceeding as in the proof of Theorem 4.1*. □

In the case of twin primes (that is, $\mathbf{a}(n) := \Lambda(n)\Lambda(n+2)$), we give an explicit description of all integers $a\geqslant -1$ (without loss of generality, since $-a(-a+2)=a(a-2)$) for which $\mu_{-1}(a,M)\neq 0$ (note the occurrence of Mersenne and Fermat primes):

a	$a(a+2)$	$\omega(a(a+2))$
-1	-1	0
1	3	1
2	8	1
$p^e, p\neq 2 : p^e+2=q^f$	p^eq^f	2
$2^e : 2^{e-1}+1=q^f$	$2^{e+1}(2^{e-1}+1)$	2
$2^e-2 : 2^{e-1}-1=q^f$	$2^{e+1}(2^{e-1}-1)$	2

Proof of Theorem 2.9. Define $\mathcal{S}:=\emptyset$ and $\mathbf{L}(x):=(\log x)^{1-\delta}$. We split the proof into two cases, depending on the size of y.

Case 1: $\log y\leqslant(\log M)^{1/2-\delta}$. The fundamental lemma of combinatorial sieve (see [6]) gives the following estimate, in the range $2\leqslant y\leqslant x^{o(1)}$:

$$\mathcal{A}(x,y) = x\prod_{p\leqslant y}\left(1-\frac{1}{p}\right)(1+E(x,y)), \tag{38}$$

where $E(x,y)\ll x^{-1/3}$ for $2\leqslant y<(\log x)^2/16$, and $E(x,y)\ll u^{-u}(\log y)^3$ for $(\log x)^2/16\leqslant y\leqslant x$, with the usual notation $u:=\log x/\log y$ (so $y^u=x$). This shows that Hypothesis 3.2 holds. One shows that

$$\mathcal{A}_d(x,y) := \sum_{\substack{n\leqslant x\\ d\mid n}}\mathbf{a}_y(n) = \begin{cases}\mathcal{A}\left(\dfrac{x}{d},y\right) & \text{if } p\mid d\Rightarrow p\geqslant y,\\ 0 & \text{else,}\end{cases}$$

so we have $\mathcal{A}_d(x,y)=\mathcal{A}(\frac{\mathbf{h}_y(d)}{d}x,y)$, where

$$\mathbf{h}_y(d) := \begin{cases}1 & \text{if } p\mid d\Rightarrow p\geqslant y,\\ 0 & \text{else.}\end{cases}$$

Wolke [20] has shown a Bombieri–Vinogradov theorem for this sequence, which states that, for any $A>0$, there exists $B=B(A)$ such that for any $Q\leqslant x^{1/2}/\log^B x$, we have, uniformly in the range $y \leqslant \sqrt{x}$,

$$\sum_{q \leqslant Q} \max_{(a,q)=1} \max_{z \leqslant x} \left| \sum_{\substack{n \leqslant z \\ n \equiv a \bmod q}} \mathbf{a}_y(n) - \frac{1}{\phi(q)} \sum_{\substack{n \leqslant z \\ (n,q)=1}} \mathbf{a}_y(n) \right| \ll \frac{x}{\log^A x}. \tag{39}$$

(Note that if $(a,q) > 1$, then $\mathcal{A}(x,y;q,a)$ is bounded.) We will only use this for $Q = 2\mathbf{L}(x)$, and so from now on we suppose that $q \leqslant 2(\log x)^{1-\delta}$. Arguing as in Section 3.1, we have, for $x/2\mathbf{L}(x) \leqslant z \leqslant x$, that

$$\begin{aligned} \frac{1}{\phi(q)} \sum_{\substack{n \leqslant z \\ (n,q)=1}} \mathbf{a}_y(n) &= \frac{1}{\phi(q)} \sum_{d \mid q} \mu(d) \mathcal{A}_d(z,y) = \frac{1}{\phi(q)} \sum_{d \mid q} \mu(d) \mathcal{A}\left(\frac{\mathbf{h}_y(d)}{d} z, y \right) \\ &= \frac{\mathcal{A}(z,y)}{\phi(q)} \sum_{d \mid q} \frac{\mathbf{h}_y(d)\mu(d)}{d} (1 + E_{d;q}(z,y)) \\ &= \frac{\mathcal{A}(z,y)}{q\gamma_y(q)} (1 + O(x^{-1/3+o(1)})), \end{aligned}$$

since by (38), in the range $d \leqslant q \leqslant (\log x)^{1-\delta}$, we have

$$E_{d;q}(z,y) \ll \left(\frac{d}{z} \right)^{1/3} \ll x^{-1/3+o(1)}.$$

Summing this over $q \leqslant 2\mathbf{L}(x)$ and using (39), we obtain that

$$\sum_{q \leqslant 2\mathbf{L}(x)} \max_{(a,q)=1} \max_{x/2\mathbf{L}(x) \leqslant z \leqslant x} \left| \mathcal{A}(z,y;q,a) - \frac{\mathcal{A}(z,y)}{q\gamma_y(q)} \right| \ll \frac{\mathcal{A}(x,y)}{\mathbf{L}(x)^{1+\delta}}. \tag{40}$$

Having a Siegel–Walfisz theorem in hand, we now prove an analogue of Proposition 5.1. A straightforward computation shows that

$$Z_{\mathcal{A}}(s) := \sum_{\substack{n \geqslant 1 \\ (n,a)=1}} \frac{1}{n^{s+1}\gamma_y(n)} = \zeta(s+1) \prod_{p \mid a} \left(1 - \frac{1}{p^{s+1}} \right) \prod_{\substack{p \nmid a \\ p < y}} \left(1 + \frac{1}{(p-1)p^{s+1}} \right) \tag{41}$$

$$= \mathfrak{S}(s)\zeta(s+1)\zeta(s+2)Z_{11}(s) \prod_{p \geqslant y} \left(1 + \frac{1}{(p-1)p^{s+1}} \right)^{-1}, \tag{42}$$

where

$$\mathfrak{S}(s) = \mathfrak{S}(s;a) := \prod_{p \mid a} \left(1 - \frac{1}{p^{s+1}} \right) \left(1 + \frac{1}{(p-1)p^{s+1}} \right)^{-1},$$

$$Z_{11}(s) = Z_{11}(s;a) := \prod_{p} \left(1 + \frac{1}{(p-1)p^{s+2}} - \frac{1}{(p-1)p^{2s+3}} \right).$$

We will now use representation (41). Representation (42) will be useful for larger values of y, since then $\prod_{p<y}(1 - 1/p^{s+2})^{-1}$ behaves like $\zeta(s+2)$ on the line $\Re(s) = -1 + 1/\log M$. Note that, by (41), $Z_{\mathcal{A}}(s)$ is defined on the whole complex plane, except at $s = 0$. As before, we need to compute the integral

$$I := \frac{1}{2\pi i} \int_{\Re(s)=2} Z_{\mathcal{A}}(s)\psi(s) \frac{ds}{s},$$

where

$$\psi(s) := \frac{(2M)^s - M^s}{s+1} + \left(\frac{x}{2M} \right)^s - \left(\frac{x}{M} \right)^s,$$

which has a double zero at $s=0$, so

$$\begin{aligned} I &= \frac{1}{2\pi i}\int_{\Re(s)=-\frac12} Z_{\mathcal{A}}(s)\psi(s)\frac{ds}{s} \\ &= \frac{1}{2\pi i}\int_{\Re(s)=-\frac12} Z_{\mathcal{A}}(s)((2M)^s-M^s)\frac{ds}{s(s+1)} + O_{a,\epsilon}\left(\left(\frac{M}{x}\right)^{1/2-\epsilon}\log y\right), \end{aligned}$$

by the same arguments as in Lemma 5.12 (and Merten's theorem). We now proceed as in the proof of Proposition 5.9. Moving the contour of integration to $\Re(s)=\sigma=-1+1/\log M$ and using the bounds $Z_{\mathcal{A}}(\sigma+it)\ll_\epsilon (|t|+1)^{1/2+\epsilon}\log y$ and $Z'_{\mathcal{A}}(\sigma+it)\ll_\epsilon (|t|+1)^{1/2+\epsilon}(\log y)^2$ for $|t|\geqslant 2$ (by Cauchy's theorem for the derivatives), we can deduce that

$$I = \frac{2\mu_y(a,M)-\mu_y(a,2M)}{2M}\left(1+O_a\left(\frac{(\log y)^2\log\log M}{\log M}\right)\right)+o(1),$$

where

$$\mu_y(a,M) := \begin{cases} -\dfrac12\displaystyle\prod_{p<y}\left(1-\frac1p\right)^{-1} & \text{if } a=\pm1,\\ 0 & \text{else.}\end{cases}$$

The condition $\log y\leqslant(\log M)^{1/2-\delta}$ ensures that the term in the O_a tends to zero. We conclude the proof in the same lines as that of Theorem 4.1*.

Case 2: $\mathbf{L}^{(1+\delta)\log\log\mathbf{L}}\leqslant y\leqslant\sqrt{x}$. Note that it is sufficient to consider this range, since $\mathbf{L}^{(1+\delta)\log\log\mathbf{L}}<(\log x)^{\log\log\log x}$. We have

$$\mathcal{A}(x,y)=\frac{x\omega(u)}{\log y}\left(1+O\left(\frac{1}{\log y}\right)\right),$$

where $u:=\log x/\log y$ and $\omega(u)$ is Buchstab's function (see [**19**, Théoréme III.6.4]). Therefore, we can use the properties of $\omega(u)$ to show that, in the range $1/\mathbf{L}(x)\leqslant z\leqslant 1+\delta$,

$$\frac{\mathcal{A}(zx,y)}{\mathcal{A}(x,y)} = z\frac{\omega(u-O(\log\mathbf{L}/\log y))}{\omega(u)}\left(1+O\left(\frac{1}{\log y}\right)\right) = z\left(1+O\left(\frac{\log\mathbf{L}}{\log y}\right)\right),$$

hence Hypothesis 3.2 holds if $y>x^{1/\log\log x}$. If $\mathbf{L}^{(1+\delta)\log\log\mathbf{L}}\leqslant y\leqslant x^{1/\log\log x}$, Hypothesis 3.2 follows from (38).

Now, since $q\leqslant 2\mathbf{L}(x)<y$, we have the equality

$$\frac{1}{\phi(q)}\sum_{\substack{n\leqslant x\\(n,q)=1}}\mathbf{a}_y(n) = \frac{1}{\phi(q)}\sum_{n\leqslant x}\mathbf{a}_y(n) = \frac{\mathcal{A}(x,y)}{q\gamma_y(q)},$$

and thus, using (39), we conclude that Hypothesis 3.1* holds. We now turn to an analogue of Proposition 5.1, which we prove using (42). We need an estimate for

$$\begin{aligned} I &= \frac{1}{2\pi i}\int_{\Re(s)=-\frac12} Z_{\mathcal{A}}(s)\psi(s)\frac{ds}{s} \\ &= \frac{1}{2\pi i}\int_{\Re(s)=-\frac12} Z_{\mathcal{A}}(s)((2M)^s-M^s)\frac{ds}{s(s+1)} + O_{a,\epsilon}\left(\left(\frac{M}{x}\right)^{1/2-\epsilon}\right), \end{aligned}$$

since on the line $\sigma=-1+1/\log M$, we have the bound

$$\prod_{p\geqslant y}\left(1+\frac{1}{(p-1)p^{s+1}}\right)^{-1}\ll\prod_{p\geqslant\mathbf{L}^{(1+\delta)\log\log\mathbf{L}}}\left(1+\frac{C_1}{p(\log p)^{1+\delta}}\right)\ll 1,$$

and similarly for the derivative of this product. We now study the function $Z(s):=Z_{\mathcal{A}}(s)/s(s+1)$. Using the bounds we just proved, we obtain that for $s=-1+1/\log M+it$

with $|t| \geqslant 2$,

$$|Z(s)|, \ |Z'(s)| \ll_\epsilon (|t|+1)^{-3/2+\epsilon}.$$

If $\omega(a) \geqslant 2$, then $Z(s)$ and $Z'(s)$ are bounded near $s=-1$ and we conclude that $I=o(1)$. If $\omega(a)=1$, then we define

$$Z_{12}(s) = Z_{12}(s;a) := Z(s) - \frac{c(M,y)}{s+1},$$

where

$$c(M,y) := -\frac{1}{2}\frac{\phi(a)}{a}\prod_{p|a}\log p \prod_{p\geqslant y}\left(1+\frac{1}{(p-1)p^{1/\log M}}\right)^{-1}.$$

One sees that, for s close to -1 with $\Re(s) = 1/\log M$,

$$\prod_{p\geqslant y}\left(1+\frac{1}{(p-1)p^{s+1}}\right)^{-1} = (1+O(|s+1|))\prod_{p\geqslant y}\left(1+\frac{1}{(p-1)p^{1/\log M}}\right)^{-1},$$

hence $|Z_{12}'(s)| \ll 1/|s+1|$, and thus

$$I = -\frac{1}{2}\frac{\phi(a)}{a}\prod_{p|a}\log p\prod_{p\geqslant y}\left(1+\frac{1}{(p-1)p^{1/\log M}}\right)^{-1}(1+o(1)). \tag{43}$$

If $a=\pm 1$, then we take $Z_{13}(s) = Z_{13}(s;a) := Z(s) - c(M,y)/(s+1)^2$, and since $Z_{13}(s) \ll 1/|s+1|$, we obtain that (43) holds. Finally, in our range of y,

$$\prod_{p\geqslant y}\left(1+\frac{1}{(p-1)p^{1/\log M}}\right)^{-1} = 1 + O\left(\frac{1}{\log y}\right). \qquad \square$$

Appendix A. Generalities on binary quadratic forms

In this section, we review several classical facts about the distribution of positive definite binary quadratic forms $Q(x,y) = \alpha x^2 + \beta xy + \gamma y^2$ in arithmetic progressions. We recall the notation $d = \beta^2 - 4\alpha\gamma$, $\mathcal{S} = \{p \mid 2d\}$, $\chi_d = (\frac{4d}{\cdot})$ and

$$R_a(q) = \#\{1 \leqslant x, y \leqslant q : Q(x,y) \equiv a \bmod q\}.$$

LEMMA A.1. *The function $R_a(q)$ is multiplicative as a function of q.*

Proof. Define $S_a(q) := \{(x,y) \in (\mathbb{Z}\cap[1,q])^2 : Q(x,y) \equiv a \bmod q\}$ and let q_1, q_2 be two coprime integers. The 'reduction mapping'

$$\begin{aligned} S_a(q_1q_2) &\longrightarrow S_a(q_1)\times S_a(q_2),\\ (x,y) \bmod q_1q_2 &\longmapsto ((x,y) \bmod q_1, (x,y) \bmod q_2)\end{aligned}$$

is a bijection by the Chinese remainder theorem. $\square$

LEMMA A.2. *Take $Q(x,y) := x^2 - dy^2$ with $d \equiv -1 \bmod 4$, and let $a \neq 0$ be a fixed integer such that $(a,2d)=1$. We have that*

$$\frac{R_a(q)}{q^2} = \frac{\mathbf{f}_a(q)}{q\gamma(q)},$$

where

$$\gamma(q) := \prod_{p\mid q}\left(1-\frac{\chi_d(p)}{p}\right)^{-1}$$

and $\mathbf{f}_a(q)$ *is a multiplicative function defined on prime powers as follows.*

For $p\nmid 2ad$, $\mathbf{f}_a(p^e):=1$. *For* $p^f \parallel a$ *with* $f\geqslant 1$ *(so* $p\nmid 2d$*),*

$$\mathbf{f}_a(p^e) := \begin{cases} e+1+\dfrac{1}{p-1} & \text{if } \chi_d(p)=1, e\leqslant f,\\ f+1 & \text{if } \chi_d(p)=1, e>f,\\ \dfrac{1}{p+1} & \text{if } \chi_d(p)=-1, e\leqslant f, 2\nmid e,\\ 1-\dfrac{1}{p+1} & \text{if } \chi_d(p)=-1, e\leqslant f, 2\mid e,\\ 0 & \text{if } \chi_d(p)=-1, e>f, 2\nmid f,\\ 1 & \text{if } \chi_d(p)=-1, e>f, 2\mid f.\end{cases} \tag{A.1}$$

For $p\mid 2d$ *(so* $p\nmid a$*),*

$$\mathbf{f}_a(p^e) = \begin{cases} 1+\left(\dfrac{a}{p}\right) & \text{if } p\neq 2,\\ 1+\left(\dfrac{-4}{a}\right) & \text{if } p=2, e\geqslant 2,\\ 1 & \text{if } p=2, e=1.\end{cases} \tag{A.2}$$

Proof. By Lemma A.1, it is enough to show that, for any prime p and integer $e\geqslant 1$,

$$\frac{R_a(p^e)}{p^e} = \frac{\mathbf{f}_a(p^e)}{\gamma(p)}. \tag{A.3}$$

First case: $p\nmid 2d$.

We will proceed as in [2, Section 2.3], by using Gauss sums. Writing $e(n):=e^{2\pi i n}$,

$$\begin{aligned} R_a(p^e) &= \frac{1}{p^e}\sum_{1\leqslant m\leqslant p^e} e\left(-m\frac{a}{p^e}\right)\left(\sum_{1\leqslant x\leqslant p^e} e\left(m\frac{x^2}{p^e}\right)\right)\left(\sum_{1\leqslant y\leqslant p^e} e\left(-md\frac{y^2}{p^e}\right)\right)\\ &= p^e + \frac{1}{p^e}\sum_{1\leqslant m\leqslant p^e-1} e\left(-m\frac{a}{p^e}\right) g(m;p^e)g(-md;p^e),\end{aligned}$$

where $g(m;q):=\sum_{n=1}^{q} e(mn^2/q)$ is a Gauss sum. We have the following properties (see [1]):

$$\text{If } q \text{ is odd, then } g(1;q)^2 = \left(\frac{-1}{q}\right)q, \tag{A.4}$$

$$\text{If } (q,m)=1, \text{ then } g(m;q) = \left(\frac{m}{q}\right)g(1;q). \tag{A.5}$$

As for Ramanujan sums (see, for example, (3.3) of [15])

$$\sum_{\substack{m=1\\(m,q)=1}}^{q} e(ma/q) = \phi(q)\frac{\mu(q/(q,a))}{\phi(q/(q,a))}. \tag{A.6}$$

Using these properties, we compute

$$\begin{aligned}
R_a(p^e) &= p^e + \frac{1}{p^e}\sum_{g=1}^{e}\sum_{\substack{1\leqslant m\leqslant p^e-1\\ p^{e-g}\|m}} e\left(-m\frac{a}{p^e}\right) g(m;p^e)g(-md;p^e)\\
&= p^e + \frac{1}{p^e}\sum_{g=1}^{e}\sum_{\substack{1\leqslant m'\leqslant p^g-1\\ p\nmid m'}} e\left(-m'\frac{a}{p^g}\right) p^{2e-2g} g(m';p^g)g(-m'd;p^g)\\
&= p^e + p^e\sum_{g=1}^{e}\left(\frac{d}{p^g}\right)p^{-g}\sum_{\substack{1\leqslant m'\leqslant p^g-1\\ p\nmid m'}} e\left(-m'\frac{a}{p^g}\right) \quad \text{by (A.4) and (A.5)}\\
&= p^e + p^e\sum_{g=1}^{e}\left(\frac{d}{p}\right)^g\left(1-\frac{1}{p}\right)\frac{\mu(p^g/(p^g,a))}{\phi(p^g/(p^g,a))} \quad \text{by (A.6)},
\end{aligned}$$

which shows (after a straightforward computation) that (A.3) holds for $p \nmid 2d$.

Second case: $p \mid 2d$, $p \neq 2$. In this case, we have that $p \nmid a$, since $(a, \mathcal{S}) = 1$. The number of solutions of $x^2 - dy^2 \equiv a \bmod p$ is exactly $p(1 + (\frac{a}{p}))$. Moreover, such a solution must satisfy $x \not\equiv 0 \bmod p$; thus by Hensel's lemma we obtain that

$$\frac{R_a(p^e)}{p^e} = 1 + \left(\frac{a}{p}\right).$$

Third case: $p = 2$. In this case, $2 \nmid a$. We have that $R_a(2) = 2$. Reducing the equation $x^2 - dy^2 \equiv a \bmod 2^e$ (using that $d \equiv -1 \bmod 4$), we obtain

$$x \not\equiv y \bmod 2, \quad x^2 + y^2 \equiv a \bmod 4,$$

which shows that there are no solutions if $a \equiv 3 \bmod 4$. Suppose now that $a \equiv 1 \bmod 4$. For $e \geqslant 3$, an odd integer is a square mod 2^e if and only if it is congruent to 1 mod 8; in fact we have the following isomorphism:

$$(\mathbb{Z}/2^e\mathbb{Z})^\times \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2^{e-2}\mathbb{Z}.$$

Using these well-known facts, we find that the number of solutions to $x^2 - dy^2 \equiv a \bmod 2^e$ such that x is odd is

$$\begin{aligned}
&= 4\#\{y \bmod 2^e : dy^2 + a \equiv 1 \bmod 8\}\\
&= 2^{e-1}\#\{y \bmod 8 : y^2 \equiv d^{-1}(1-a) \bmod 8\} = 2^e
\end{aligned}$$

since $d^{-1}(1-a) \equiv 0, 4 \bmod 8$. Now the number of solutions of $x^2 - dy^2 \equiv a \bmod 2^e$ such that x is even is just the number of solutions of $y^2 - d^{-1}x^2 \equiv -d^{-1}a \bmod 2^e$ such that y is odd, which as we have shown (and using that $-d^{-1} \equiv 1 \bmod 4$) is equal to 2^e. We conclude that

$$\frac{R_a(2^e)}{2^e} = \begin{cases} 2 & \text{if } a \equiv 1 \bmod 4,\\ 0 & \text{if } a \equiv 3 \bmod 4.\end{cases}$$

$\square$

Lemma A.3. *Take* $Q(x,y) := \alpha x^2 + \beta xy + \gamma y^2$ *with* $(\alpha, \beta, \gamma) = 1$ *and* $d = \beta^2 - 4\alpha\gamma \equiv 1, 5, 9, 12, 13 \bmod 16$. *Let* $a \neq 0$ *be a fixed integer with* $(a, 2d) = 1$. *We have, for* $(q, 2d) = 1$, *that*

$$\frac{R_a(q)}{q^2} = \frac{\mathbf{f}_a(q)}{q\gamma(q)},$$

where

$$\gamma(q) := \prod_{p \mid q} \left(1 - \frac{\chi_d(p)}{p}\right)^{-1}$$

and $\mathbf{f}_a(q)$ *is defined as in Lemma* A.2. *Moreover, for* $p \mid 2d$, $p \neq 2$ (*so* $p \nmid a$),

$$\frac{R_a(p^e)}{p^e} = \begin{cases} 1 + \left(\frac{\alpha a}{p}\right) & \text{if } p \neq 2, p \nmid \alpha, \\ 1 + \left(\frac{\gamma a}{p}\right) & \text{if } p \neq 2, p \nmid \gamma, \end{cases} \tag{A.7}$$

and

$$\frac{R_a(2^e)}{2^e} = \begin{cases} 1 & \text{if } 2 \mid \beta, e = 1, \\ 1 + \left(\frac{-4}{\alpha a}\right) & \text{if } 2 \mid \beta, 2 \nmid \alpha, e \geqslant 2, \\ 1 + \left(\frac{-4}{\gamma a}\right) & \text{if } 2 \mid \beta, 2 \nmid \gamma, e \geqslant 2, \\ \frac{1}{2} & \text{if } 2 \nmid \beta, 2 \mid \alpha\gamma, \\ \frac{3}{2} & \text{if } 2 \nmid \alpha\beta\gamma. \end{cases} \tag{A.8}$$

Proof. First write $Q(x,y)$ in four different ways:

$$Q(x,y) = \frac{1}{4\alpha}((2\alpha x + \beta y)^2 - dy^2) \tag{A.9}$$

$$= \frac{1}{\alpha}\left(\left(\alpha x + \frac{\beta}{2}y\right)^2 - \frac{d}{4}y^2\right) \tag{A.10}$$

$$= \frac{1}{4\gamma}((\beta x + 2\gamma y)^2 - dx^2) \tag{A.11}$$

$$= \frac{1}{\gamma}\left(\left(\gamma y + \frac{\beta}{2}x\right)^2 - \frac{d}{4}x^2\right). \tag{A.12}$$

We will split into five distinct cases.

Case 1 : $p \nmid 2\alpha$. In this case, we use the representation (A.9). Note that the mapping $\phi_y : x \mapsto 2\alpha x + \beta y$ is an automorphism of $\mathbb{Z}/p^e\mathbb{Z}$, so

$$R_a(p^e) = \#\{1 \leqslant x, y \leqslant p^e : x^2 - dy^2 \equiv 4\alpha a \bmod p^e\}.$$

Going through the proof of Lemma A.2, we see that

$$\frac{R_a(p^e)}{p^e} = \frac{\mathbf{f}_{4\alpha a}(p^e)}{\gamma(p)} = \frac{\mathbf{f}_{\alpha a}(p^e)}{\gamma(p)} \qquad \left(= \frac{\mathbf{f}_a(p^e)}{\gamma(p)} \text{ if } p \nmid d\right).$$

Case 2 : $p \nmid 2\gamma$. In this case, we proceed in an analogous way to the first case, using the representation (A.11) to obtain that

$$\frac{R_a(p^e)}{p^e} = \frac{\mathbf{f}_{4\gamma a}(p^e)}{\gamma(p)} = \frac{\mathbf{f}_{\gamma a}(p^e)}{\gamma(p)} \qquad \left(= \frac{\mathbf{f}_a(p^e)}{\gamma(p)} \text{ if } p \nmid d\right).$$

Case 3 : $p \mid \alpha$, $p \mid \gamma$, $p \neq 2$. In this case $p \nmid \beta$, so $p \nmid d$. Writing $X := x + y$ and $Y := y$, we compute that

$$\alpha X^2 + \beta XY + \gamma Y^2 = \alpha x^2 + (2\alpha + \beta)xy + (\alpha + \beta + \gamma)y^2 =: \alpha' x^2 + \beta' xy + \gamma' y^2.$$

We have $p \mid \alpha'$, $p \nmid \beta'$ and $p \nmid \gamma'$, which reduces the problem to Case 2, and so

$$\frac{R_a(p^e)}{p^e} = \frac{\mathbf{f}_{(\alpha+\beta+\gamma)a}(p^e)}{\gamma(p)} = \frac{\mathbf{f}_a(p^e)}{\gamma(p)}.$$

Case 4.1 : $p = 2$, $2 \mid \beta$. In this case, $d \equiv 0 \bmod 4$. We have that either $2 \nmid \alpha$, or $2 \nmid \gamma$. In the first event we use representation (A.10), which gives

$$R_a(2^e) = \#\{1 \leqslant x, y \leqslant 2^e : x^2 - d'y^2 \equiv \alpha a \bmod 2^e\}$$

with $d' := \frac{d}{4} \equiv -1 \bmod 4$. Going back to the proof of Lemma A.2, we obtain that

$$\frac{R_a(2^e)}{2^e} = \mathbf{f}_{\alpha a}(2^e).$$

In the event that $2 \nmid \gamma$, the result is

$$\frac{R_a(2^e)}{2^e} = \mathbf{f}_{\gamma a}(2^e).$$

Note that if $2 \nmid \alpha\gamma$, then, since $\frac{d}{4} \equiv -1 \bmod 4$, we have $\alpha \equiv \gamma \bmod 4$, so

$$\mathbf{f}_{\alpha a}(2^e) = \mathbf{f}_{\gamma a}(2^e).$$

Case 4.2 : $p = 2$, $2 \nmid \beta$. In this case, $2 \nmid d$ and $2 \nmid a$. An easy application of Hensel's lemma in either of the variables x or y (since one of them has to be odd) yields

$$\frac{R_a(2^e)}{2^e} = \frac{R_a(2)}{2},$$

and all the possibilities are contained in the following table.

$\alpha \bmod 2$	$\beta \bmod 2$	$\gamma \bmod 2$	$R_a(2)$
0	1	0	1
0	1	1	1
1	1	0	1
1	1	1	3

$\square$

Acknowledgements. I would like to thank my supervisor Andrew Granville for suggesting this generalization and for his help and advice in general. I would also like to thank my colleagues Farzad Aryan, Mohammad Bardestani, Dimitri Dias, Tristan Freiberg and Kevin Henriot for many fruitful conversations. Finally, I thank the anonymous referee for his careful reading of the paper and his helpful comments. Ce travail a été rendu possible grâce à des bourses doctorales du Conseil de Recherche en Sciences Naturelles et en Génie du Canada et de la Faculté des Études Supérieures et Postdoctorales de l'Université de Montréal.

References

1. B. C. Berndt, R. J. Evans and K. S. Williams, *Gauss and Jacobi sums*, Canadian Mathematical Society Series of Monographs and Advanced Texts, A Wiley-Interscience Publication (Wiley, New York, 1998) xii+583 pp.
2. V. Blomer, J. Brüdern and R. Dietmann, 'Sums of smooth squares', *Compos. Math.* 145 (2009) 1401–1441.
3. E. Bombieri, J. B. Friedlander and H. Iwaniec, 'Primes in arithmetic progressions to large moduli', *Acta Math.* 156 (1986) 203–251.

4. E. Bombieri, J. B. Friedlander and H. Iwaniec, 'Primes in arithmetic progressions to large moduli. II', *Math. Ann.* 277 (1987) 361–393.
5. E. Bombieri, J. B. Friedlander and H. Iwaniec, 'Primes in arithmetic progressions to large moduli. III.' *J. Amer. Math. Soc.* 2 (1989) 215–224.
6. N. Govert De Bruijn, 'On the number of uncancelled elements in the sieve of Eratosthenes', *Nederl. Akad. Wetensch. Proc.* 53 (1950) 803–812.
7. H. Davenport, *Multiplicative number theory*, 3rd edn, Revised and with a preface by Hugh L. Montgomery, Graduate Texts in Mathematics 74 (Springer, New York, 2000) xiv+177 pp.
8. H. M. Edwards, *Riemann's zeta function*. Pure and Applied Mathematics 58 (Academic Press, New York/London, 1974) xiii+315 pp.
9. D. Fiorilli, 'Residue classes containing an unexpected number of primes', *Duke Math. Journal*, to appear.
10. J. B. Friedlander and A. Granville, 'Relevance of the residue class to the abundance of primes', *Proceedings of the Amalfi Conference on Analytic Number Theory (Maiori, 1989)*(University of Salerno, Salerno, 1992) 95–103.
11. A. Granville and K. Soundararajan, 'An uncertainty principle for arithmetic sequences', *Ann. of Math.* (2) 165 (2007) 593–635.
12. C. Hooley, 'On the Barban–Davenport–Halberstam theorem. I. Collection of articles dedicated to Helmut Hasse on his seventy-fifth birthday, III', *J. Reine Angew. Math.* 274/275 (1975) 206–223.
13. M. N. Huxley, 'Exponential sums and the Riemann zeta function. V', *Proc. London Math. Soc.* (3) 90 (2005) 1–41.
14. H. Iwaniec, 'The half dimensional sieve', *Acta Arith.* 29 (1976) 69–95.
15. H. Iwaniec and E. Kowalski, *Analytic number theory*. American Mathematical Society Colloquium Publications 53 (American Mathematical Society, Providence, RI, 2004) xii+615 pp.
16. K. Kawada, The prime k-tuplets in arithmetic progressions, *Tsukuba J. Math.* 17 (1993) 43–57.
17. V. A. Plaksin, 'Asymptotic formula for the number of solutions of an equation with primes', *Izv. Akad. Nauk SSSR Ser. Mat.* 45 (1981) 321–397.
18. V. A. Plaksin, 'Asymptotic formula for the number of representations of a natural number by a pair of quadratic forms, the arguments of one of which are primes', *Izv. Akad. Nauk SSSR Ser. Mat.* 48 (1984) 1245–1265.
19. G. Tenenbaum, *Introduction à la théorie analytique et probabiliste des nombres*, Deuxième édition, Cours Spécialisés 1 (Société Mathématique de France, Paris, 1995) xv+457 pp.
20. D. Wolke, 'Über die mittlere Verteilung der Werte zahlentheoretischer Funktionen auf Restklassen. II', *Math. Ann.* 204 (1973) 145–153.

Daniel Fiorilli
Département de mathématiques et de statistique
Université de Montréal
CP 6128, succ. Centre-ville
Montréal, QC
Canada H3C 3J7

Current address:
Department of Mathematics
University of Michigan East Hall
530 Church Street
Ann Arbor, MI 48109-1043
USA

fiorilli@umich.edu